



Synthesis of Multi-Modal Socially Intelligent Human-Robot Interaction

Karen Tatarian

► To cite this version:

Karen Tatarian. Synthesis of Multi-Modal Socially Intelligent Human-Robot Interaction. Artificial Intelligence [cs.AI]. Sorbonne Université, 2022. English. NNT : 2022SORUS041 . tel-03863734

HAL Id: tel-03863734

<https://theses.hal.science/tel-03863734>

Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PhD Thesis

This dissertation is submitted for the degree of

Docteur de Sorbonne Université

Ecole Doctorale Sciences Mécaniques, Acoustique, Electronique et Robotique (SMAER)

Karen Tatarian

Synthesis of Multi-Modal Socially Intelligent Human-Robot Interaction

Supervised by Professor Dr. Mohamed Chetouani

submitted in December 2021

Thesis jury:

Prof. Dr. Jonathan Gratch	Reviewer	University of Southern California
Prof. Dr. Dirk Heylen	Reviewer	University of Twente
Prof. Dr. Amal El Fallah Seghrouchni	Examiner	Sorbonne University
Dr. Joost Broekens	Examiner	Leiden University
Prof. Dr. Mohamed Chetouani	Supervisor	Sorbonne University



Dedication

To my beautiful parents Paul & Rita, who have always encouraged me to reach my full potential and supported me throughout the years .. To my grandpa, General Georges, who always saw so much greatness in me even when I could not .. To the endlessly curious little girl I was and still am, who found so much joy and wonder in mathematics and physics
... this one is for you

"....We must have perseverance and above all confidence in ourselves. We must believe that we are gifted for something, and that this thing, at whatever cost, must be attained."

- Marie Skłodowska-Curie

Acknowledgements

This thesis is the fruit of 3 years of research and work with Professor Mohamed Chetouani at Sorbonne Université and with the support of Marine Chamoux from SoftBank Robotics Europe (SBRE). I would also like to thank the entire Expressivity Team at SBRE for their support throughout my PhD.

I want to also acknowledge the Marie Skłodowska-Curie Fellowship through the ANI-MATAS project as this thesis received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 765955.

Abstract

By merely observing humans, one can directly infer that no social interaction takes place without cues, whether verbal or nonverbal, that allow others to interpret behaviors and reasonably estimate intentions. These powerful social signals and nonverbal behaviors are complex and multi-modal, which means they are made of different combinations of modalities and cues like gestures, gaze behavior, and proxemics (e.g., management of space and environment). Thus for a robot to be perceived as a socially-intelligent agent by humans, it is expected to be able to hold a successful social interaction, adapt to the social environment, and exhibit appropriate multi-modal behavior. This thesis first investigates how one of these modalities can help adapt another one, then explores the effects of the modalities when performed multi-modally on behavioral interaction outcomes and perception of the robot's social intelligence, and finally presents an architecture using reinforcement learning for the robot to learn to combine its multi-modal behaviors with a reward function based on the multi-modal social signals of the human in an interaction.

Modalities are coupled in nature and for one to adapt to the changes in the environment it may need to rely on other sensory modalities. For instance, in the first work, the robot would autonomously adapt its gaze pattern to social interaction changes, i.e groups being formed around it, based on its proxemics, which was used to estimate the roles of participants in group formations around the robot, such as active speaker, bystander, or overhearer. A pilot study looked into group formations made of a robot and two users. Results showed that participants stood closer to the adaptive robot and ranked it higher in perceived adaptability and perceived sociability as well as feeling attended to by the robot in comparison to a robot that switched its gaze attention based on new sensory detection. In the second work, multi-modal behaviors made up of gaze mechanisms, which are turn-taking, turn-yielding, floor-holding, and joint attention, social gestures, which are emblem, deictic, and beat gestures, proxemics, through social navigation, and social dialogue were implemented autonomously and studied by extracting one modality in each condition and looking into behavioral outcomes and subjective measurements. The data collection included 105 participants in a seven minutes interaction alone with the robot to investigate behavioral outcomes including but are not limited to distances of the users, speaking time, greetings performed, as well as backchannels. The study showed the extent of which each modality within the multi-modal

behavior allows the robot to influence how close the human stands, how they address the robot, whether they take its suggestions or not, and how they greet and end the interaction by mirroring nonverbal behaviors of the robot. Finally and after diving into a review of reinforcement learning for adaptation in social robotics presented in the thesis, the last work looks into the multi-modal social signals of the human, including distances, time spent looking and deciding, as well as decisions made, to formulate the reward signal, which also includes cost functions for the complexity of the multi-modal behavior performed by the robot and whether or not the human followed the robot's recommendation. This reward function was then used to adapt the robot's multi-modal behavior creating various possible combinations, which are made of gaze, gestures, proxemics, and emotional expressions, with the goal to increase the robot's social intelligence and influence. In conclusion, this thesis work dives into understanding how nonverbal modalities forming multi-modal behavior shape the success of interactions as well human behaviors and attitudes about and towards the robot and thus paving the way for a learning architecture allowing the robot to further adapt to human preferences and social signals. These findings are crucial to unlock the advancement of the social intelligence of future technologies in adapting to humans, learning from them, and communicating with them beyond just verbal means.

Table of contents

List of figures	xv
List of tables	xix
1 Introduction	1
1.1 Motivation	1
1.1.1 State-of-Art	2
1.1.2 Limitations & Challenges	3
1.1.3 Research Questions and Contributions	4
1.2 Outline of Thesis	6
1.2.1 Part I: Related Work	6
1.2.2 Part II: Contributions	6
I Related Work	9
2 State-of-the-Art	11
2.1 Proxemics	11
2.2 Gaze	13
2.3 Gestures	15
2.4 Role Emotional Expressions in Decision Making	16
2.5 Multi-Modal Social Behaviors	17
2.6 Behavior and Social Intelligence and Influence	17
3 Review on Reinforcement Learning in Social Robotics for Adaptation	20
3.1 Reinforcement Learning	20
3.1.1 Multi-Armed Bandit	22
3.1.2 Q-Learning	22
3.1.3 Background and Definitions	23

3.2	Reinforcement Learning In Social Robotics	24
3.2.1	RL for Robot Adaptation and Selecting Appropriate Behaviors	25
3.2.2	Exploring Reward Functions for RL in HRI	31
II	Contributions	38
4	Robot Gaze Adaptation using Human Proxemics	40
4.1	Introduction	41
4.2	Background and Definitions	42
4.3	Adaptive System	42
4.3.1	Conversational Role Coordinator Using Human Proxemics	43
4.3.2	Adaptive Robot Gaze Behavior	47
4.4	Study Design	47
4.4.1	Conditions	48
4.4.2	Hypotheses	49
4.4.3	Measurements	49
4.4.4	Experimental Procedure	50
4.5	Results	50
4.6	Discussion	53
4.7	Conclusion	55
5	Investigating the Synthesis & Effects of Multi-Modal Robot Behavior	56
5.1	Introduction	57
5.2	Implemented System for Multi-Modal Behavior	58
5.2.1	Social Gaze Mechanisms	59
5.2.2	Gestures	62
5.2.3	Proxemics for Approaching Human	62
5.3	Design Method and Evaluation	63
5.3.1	Social Dialogue	64
5.3.2	Design & Materials	65
5.3.3	Participants & Procedure	65
5.3.4	Hypotheses	67
5.4	Results	68
5.4.1	Distances from Pepper	68
5.4.2	Self-Disclosure	70
5.4.3	Accepting Water Offered	71
5.4.4	Social Behavior	72

5.4.5	Liking	77
5.4.6	Voice Recognition Errors	80
5.5	Discussion	80
5.6	Conclusion	83
6	Adaptation of Multi-Modal Behaviors using Reinforcement Learning	85
6.1	Introduction	86
6.2	RL Model	87
6.2.1	State-Space	87
6.2.2	Action-Space	88
6.2.3	Reward Function	89
6.3	Evaluation	90
6.3.1	Experiment Study	90
6.3.2	Simulation of Human and Interaction	91
6.4	Implementation of Reinforcement Learning Models	94
6.4.1	Multi-Armed bandit	95
6.4.2	Q-Learning	95
6.5	Results	96
6.5.1	Multi-Armed Bandit	96
6.5.2	Q-Learning	103
6.6	Discussion	107
6.7	Limitations	110
6.8	Conclusion	110
7	Discussion & Conclusion	111
7.1	Discussion	111
7.1.1	Thesis Overview & Contributions	111
7.1.2	Multi-Modal Behaviors & Social Intelligence: Perspective	115
7.1.3	Limitations & Future Work	117
7.2	Conclusion	117
	References	118

List of figures

3.1	Standard Reinforcement Learning framework (inspired by [161])	21
3.2	Map of Reinforcement Learning algorithms (inspired by [191])	22
4.1	Set-up of Transactional, Front, and Gaze Zones	43
4.2	Representation of Human Face Frame, Base Frame of Robot, and Gaze Frame of Robot	44
4.3	Flow Chart of the adaptive gaze behavior mechanism based on the status evaluation of the conversational role coordinator	45
4.4	Overview of two sets of interactions between the robot and participants forming a small group. The statuses represent each conversational role for each participant estimated by robot in different contexts and times during the interaction	48
4.5	Distances (meters) maintained by participants per condition and per role they played during the interaction: when they were playing a bystander role, main player role, and one of the two players. There was a (**) significance, for $p < 0.01$, between the two conditions. In addition, (*) represents significance of $p < 0.05$ between the two conditions for the Bystander Player and (***) represents significance of $p < 0.0001$ between different roles in the group interaction.	51
4.6	α Angles (degree) maintained by participants per condition and per role they played during the interaction: when they were playing a bystander role, main player role, and one of the two players. Alpha angle refers to the angle define in Figure 4.1	52
4.7	Ratings of ALMERE Constructs per condition by participants with (*) significance for $p < 0.05$ between the two condition. (**) refers to $p < 0.01$ and (***) to $p < 0.001$	53
4.8	Ratings from extra questions related to 'Feeling Attended To' per condition by participants with (*) ($p < 0.05$) significance between the two conditions . .	53

5.1	Summary of Multi-Modal Social Cues	58
5.2	Sample of time-line including speech, gaze mechanisms (turn-taking, floor-holding, turn-yielding), and social gestures (deictic gestures: "You" vs "Me" if mentioned in speech, beat gesture: emphasizing the two choices user needs to select from)	59
5.3	Summary of gaze mechanisms	60
5.4	Turn Yield Gaze Mechanism	61
5.5	Summary of flow of interaction	63
5.6	Schematics of the experimental room set-up with the robot during the travel agent scenario	66
5.7	Initial Distance (meters) by participants per condition at the beginning of the interaction with the robot before proxemics	69
5.8	Maintained Distance (meters) by participants per condition throughout the interaction with the robot after proxemics.	70
5.9	Total speaking time (seconds) of participants per condition during self-disclosure open-ended question	71
5.10	Number of times the water offer was accepted by participants per condition. <i>Present</i> refers to appearance of the participant's behavior of accepting the water offered by the robot by grabbing the cup and/or drinking the water, while <i>Absent</i> refers to lack of this behavior and as such not accepting the water offered by the robot.	72
5.11	Number of <i>greeting</i> waves performed by participants at the beginning of the interaction (with 95% CI errors). <i>Present</i> refers to appearance of the participant's behavior of waving to the robot at the beginning of the interaction, while <i>Absent</i> refers to lack of this behavior, e.g., the participants not waving to the robot.[*] significant at the $p < .05$ level. [***] significant at the $p < .001$ level	73
5.12	Number of <i>closing</i> waves performed by participants at the end of the interaction(with 95% CI errors). <i>Present</i> refers to appearance of the participant's behavior of waving to the robot at the end of the interaction, while <i>Absent</i> refers to lack of this behavior, e.g., the participants not waving to the robot.[*] significant at the $p < .05$ level	75
5.13	Number of back-channels performed while interacting with the robot in each condition (with 95% CI errors). <i>Present</i> refers to appearance of the back-channels performed by the participants, while <i>Absent</i> refers to lack of back-channels detected during the interaction.[*] significant at the $p < .05$ level. [**] significant at the $p < .01$ level	76

5.14	Mean Liking ranked by participants for the robot in each condition	78
5.15	Number of addressee terms used to address the robot in each condition. The addressee terms are placed in four categories: absent, no term was used to directly address the robot, informal/friendly pronoun, formal/acquaintance pronoun, and name of the robot	79
5.16	Number of Utterances while using the tablet of the robot in each condition. <i>Present</i> refers to detection of utterances/speech while tablet was used by the participants, while <i>Absent</i> refers to lack of detection.	80
6.1	Overview of the multi-modal adaptive behavior system	87
6.2	Schematics of the experimental room set-up with the robot of data-set used (Image from Tatarian et al. [168])	91
6.3	Histograms of data distributions for the simulation. Figure(a) shows the normal distributions of the distances held by the user dh (meters) when the robot's behavior includes proxemics and when the robot's behavior does not include proxemics. Figure(b) shows the normal distributions of the time spent looking at the referred object when the user takes the recommendation of the robot and when the use does not. Figure(c) shows the normal distributions of the time the human spent making a decision when the robot's behavior includes gaze and/or gesture and when the robot's behavior does not include gaze nor gesture.	93
6.4	The learning curves showing the episodic average social signal reward per algorithm after 100 (left) and 5000 (right) iterations per 20 random seeds episodes respectively. The mean and standard deviation over the 20 random seeds are plotted. The reward function used contained only the human's social signals part.	97
6.5	Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 100 per 20 episodes with the social signal reward respectively	98
6.6	Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 5000 iterations per 20 episodes with the social signal reward respectively	99
6.7	Average episodic reward per algorithm after (a) 100 (b) 1,000 and (c) 10000 iterations for 20 random seeds episodes with the full reward respectively($\delta = 0.24$). The mean and standard deviations are plotted.	100
6.8	Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after (a) 100 and (b) 10000 iterations per 20 episodes with the full reward respectively($\delta = 0.24$). The average percentage of the 20 random seeds is shown.	101

6.9	The percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 10000 iterations per 20 episodes with the full reward respectively($\delta = 0.24$). The average percentage of the 20 random seeds is shown.	101
6.10	Average reward per episode of algorithms using full reward with weight $\delta = 0.09$ after 10000 iterations for 20 random seed episodes	102
6.11	Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm using full reward with weight $\delta = 0.09$ after 10000 iterations for 20 random seed episodes	102
6.12	Average episode reward after 15000 episodes for 10 random seeds. The means and standard deviations are plotted. ϵ -MAB($\epsilon = 0.1, \alpha = 0.2$) serves as a baseline.	103
6.13	Q-Table with the average $q(s, a)$ values after 15000 episodes for 10 random seeds for Q-Learning(ϵ -decay) algorithm.. . . .	105
6.14	Average percentage of states visited after 15000 episodes for 10 random seeds for Q-Learning (ϵ -decay) algorithm.	106
6.15	Average Simulated Human's Distance(meters), Average Times Human Took Suggestion of Robot, Average Simulated Human's Time Looking at Object (ms), and Average Simulated Human's Time Deciding (ms), after 15000 episodes of 10 random seeds, by each algorithm: ϵ -MAB($\epsilon = 0.1, \alpha = 0.2$) and Q-Learning(ϵ -decay). The mean and standard deviations are plotted.	107

List of tables

3.1	Summary of Adaptation Model using Thresholding in Tanveska et al.(2020) .	25
3.2	Summary of RL Model in Bagheri et al.(2020)	26
3.3	Summary of RL Model in Hussein et al.(2017)	27
3.4	Summary of Meta-RL adaptation for trust in Gao, Sibirtseva, et al.(2019) . .	29
3.5	Summary of RL Model "TAMER" by Knox et al. (2013)	30
3.6	Summary of Proposed Simulated RL Model by Ritschel et al. (2018) [17] . .	32
3.7	Summary of Weber et al. (2018) and Ritschel et al. (2020) for their approach in using RL to have an adaptive multimodal joke generator for a social robot	34
3.8	Summary of Proposed Empathy RL Model by Leite et al. (2012) [105]	36
4.1	Extra Questions: Feeling Attended To (cronbach's alpha = 0.71)	49
4.2	Summary of average distances held by participants throughout the different roles in the interaction per condition and in total	51
5.1	Social dialogue designs per condition	64
5.2	Likeability Scale ranked from 1 to 7	66
5.3	Kruskal-Wallis Tests for the Effect of Condition on Initial and Maintained Distances	69
5.4	Comparison of the likelihood of producing opening and closing waves in each condition	74
5.5	Number of greeting and/or closing turns, which may be verbal or non-verbal, present and absent in each modality in beginning and end of the interaction. The rows represent the statuses in each condition as follows: 1) participants who greeted the robot and also closed the interaction (no significance) 2) those who greeted the robot but did not close the interaction 3) those who did not greet the robot but closed the interaction 4) those who did not greet the robot nor did they close the interaction	75
5.6	Comparison of the likelihood of producing back-channels in each condition . .	76

5.7	Frequency of modes of address used by participants towards Pepper in each condition	78
5.8	Comparison between modes of address used by participants towards Pepper in each condition	79
6.1	$k = 15$ representing $[0, k[$ behaviors of varying degrees of modality	94

Chapter 1

Introduction

When Thorndike [169] first introduced the concept of social intelligence back in 1920 by observing human-human interactions, the definition was reduced to the following: social intelligence is the ability to understand and manage others and to "act wisely in human relations". Thorndike also added that social intelligence can be best measured during social interactions since it requires human beings to respond and adapt the responses, face, voice, and gestures accordingly. A few decades and since the beginning of its origin, artificial intelligence (AI) has been heavily driven and influenced by human intelligence. So it is no surprise that when robots started gaining attention, the notion of building human-like embodied robots and machines put forward the need to synthesize a human-like intelligence to allow the robot to interact with the environment around it and to be expressive [30, 29]. This brought forward the desire to achieve social intelligence for robots to make them smarter, have more believable behaviors, and better equipped in adaptation to social situations and interactions [40]. Moreover, this has become crucial as more robots and agents are found in human-centered fields such as healthcare, service assistance, retail, and even homes. Thus, the need to address and solve for socially intelligent human-robot interactions has become central to their success and evolution.

1.1 Motivation

Social Intelligence is defined to include two aspects that need to be considered: behavioral [57], which refers to the ability to communicate and hold successful interaction, and a cognitive [18], which refers to the ability to socially adapt to the other members of the interaction and varying situations; while still considering social intelligence a multi-faced construct [98, 153]. Moreover, human behavior is naturally multi-modal, where gaze aversion mechanisms, gestures, body postures and orientations, and facial expressions play functional roles

in regulating interactions and communicating intention and engagement [86, 65]. Moreover, in his quest to formulate the success of an interaction, Erving Goffman highlights that this success depends on the situation the person is in and the expressive order, which regulates the flow of events through a specific order of social actions [61]. In parallel, a robot is an embodied agent that can navigate to adjust its position in the space, avert its gaze to portray cognitive processes, gesture using arm motions to emphasize key points, display joy or anger, among others, to influence engagement, communicate effectively, and manipulate the environment around it [11, 115, 132, 166, 12]. Thus, for a robot to be perceived as a socially-intelligent agent, it must be able to hold a successful social interaction, adapt to the social environment, and exhibit appropriate multi-modal behavior. The twofold of this bidirectional relationship between the multi-modal behaviors of the robot and that of the human it is interacting with needs to be addressed. The robot not only has to display multi-modal behavior to influence interaction outcomes with the human user, but it must also adapt to social signals and behaviors of humans. The thesis focuses on these two aspects of an interaction to present the synthesis of multi-modal socially intelligent human-robot interaction through investigating the effect of multi-modal behaviors of the robot on the outcomes of an interaction and proposing an adaptation and personalisation of these behaviors by taking into account the human social signals using Reinforcement Learning (RL).

1.1.1 State-of-Art

In the past decades, there has been considerable interest among psychologists, sociologists, cognitive scientists, and neuroscientists to investigate non-verbal behaviors observed in humans and used as communication methods and tools. Inspired by those findings, social roboticists have tried to synthesize these modalities on different robots in order to study their impact on human-robot interaction (HRI). Multi-modal behaviors are made up of several modalities carrying various functions. Some of these modalities include but are not limited to: *proxemics*, which is the study of social spatial behavior [115, 114], *vocalics*, which is the study of the nonverbal characteristics of speech such as volume, pitch, and rate [128], *oculesics*, which is the study of gaze behavior [125, 81], *kinesics*, which is the study of communicative body movement such as gestures [6, 48], and *chronemics*, which is the study of timing in social interactions [34, 123]. More recently, there has been more attention given to the modalities of *haptics*, which focuses on touch for communication by force, vibration, texture, or hand motions [187, 9] as well as *emotional expressions*, which allow the robot to simulate emotions by using facial expressions and/or body movements [132, 190]. However, while most work and studies focus on examining one modality or a combination of two at most [69, 142, 133, 184], there is a lack of comprehension of the impact and relevance

of these modalities on the social signals of the human and the interaction outcomes when performed in combination with several modalities forming multi-modal behavior. This thesis investigates and contributes to the synthesis and adaptation of the following modalities: gaze, gesture, proxemics and how they are combined to form multi-modal behaviors for robots. While this thesis does not contribute to the synthesis of emotional expressions for robots, it explores their effects when combined with other modalities to compose multi-modal behaviors.

Furthermore, a fundamental feature of achieving multi-modal socially intelligent interaction is the ability to adapt to social interactive environments [40]. While interacting with others, humans have a intrinsic tendency to adapt their manners with others by modifying their behaviors and even tones and speech to the perceived needs of their interactive partners [166, 107]. Moreover, adaptation and personalization is crucial in HRI to bring forth user-personalized and socially intelligent interactions to fit for different user profiles and cultures specially for elderly care [26, 150], assistive robots [185, 3], collaborative learning [136], and many other use cases. While some have attempted to solve adaptation by creating cognitive architecture for autonomous behavior, this approach still requires more understanding on the perception of humans and their intentions, their past memories, and anticipation of future behavior [166]. On the other hand, in reinforcement learning (RL), the agent learns through trial-and-error by interacting with the environment around it much like how one would expect social robot to behave. As such RL may provide a possible solution for achieving adaptation in HRI as a way for the robot to evaluate its behavior. While RL for adaptation in HRI is shown to be a promising approach [119, 58], it is still relatively new and not much work have been done on it. This thesis explores a RL model for adaptation in HRI by also providing potential solutions to current gaps and limitations.

1.1.2 Limitations & Challenges

First, while several studies have combined two or more of the robot’s modalities for specific functionalities in order to assess the interaction between them, there is not enough work done on studying larger sets of multi-modal behaviors. For instance, the most common studies have focused on gaze behaviours combined with gestures [69] or proxemics [184, 54]. Work combining other modalities, or comparing more than two or three modalities is rare. There is also a large body of work comparing behaviors on robots categorized as ‘social’ versus ‘non-social’ (present versus non-present) [181, 130]. However, the manipulations in these cases typically combine multiple modalities and evaluate the overall system performance, as opposed to investigating the effect of specific modalities. In addition, even when looking at individual modalities of the robot’s behaviors, there has been mixed results and findings. For instance, some studies show a positive effect of gesture on outcomes such as sociability,

competence, engagement and enjoyment [20, 90, 143, 8], whilst others show no effect on trust, persuasiveness, perceived human-likeness, or rapport [164, 160, 69]. Similar mixed findings were found in exploring gaze behavior in robotics. In some studies, robots using gaze cues have been found to improve impressions of competence and intelligence [125, 78], however, in others it did not [159, 11]. As such, although potentially useful for coordinating task based interactions and establishing participant roles, the importance of gaze in developing more subjective robot impressions is yet to be determined. Conclusively, there is a clear need to develop a more comprehensive perspective on how different combinations of modalities, specially gaze, gesture, and proxemics, contribute to overall perceptions of social intelligence during the course of an interaction.

Second, RL for the robot’s capabilities and learning from nonverbal social cues from the human during HRI is still to a great extent underdeveloped. This has been mostly due to two main reasons and limitations shown in the recent studies [89]: first due the unpredictable and varying multi-modal human social signals and behaviors and second due to the high-dimensional continuous state-space this creates for the robot to interact in such scenarios. While some have tried to discretize the social signals to form smaller state spaces [88, 17, 58, 96, 105], this may lead to a loss of information as well as prevent more generalized models to deal with task-independent and various scenarios. In addition, there is a lack of studies addressing the robot learning how to adapt multi-modal behaviors during HRI for task-independent scenarios. Moreover, training of RL for HRI can be very costly and challenging specially during the current COVID-19 pandemic. As such, contributions for simulation of HRI to train RL algorithms for social robots are called for. Finally, there is a need to study a generalized RL model for the robot to learn how to combine multi-modal behaviors to adapt and personalize to the multi-modal social signals of the human during HRI.

1.1.3 Research Questions and Contributions

The main research question this thesis aims at contributing to is the following: *"When found in an interaction set-up, how should the robot use its multi-modal behavior and understanding of the interaction dynamics with the humans around it to act as a socially intelligent agent?"*. The thesis focuses on trying to answer the question by exploring three fundamental sub-questions: How can a human behavior be used to allow the robot to adapt to varying interaction setups? How does each modality impact the outcomes of an interaction specially when it is performed in a multi-modal manner by the robot and in which aspect of the interaction is each modality relevant and significant? The findings of these questions would help answer the third research question: how can the robot use human social signals during

an interaction to adapt and personalize its own multi-modal behavior? Thus exploring the multi-modal combinations formed as the robot interacts with the user.

First, group interactions were employed as a use case to contribute to the first question. The thesis provides an adaptive autonomous rule-based system that allows the robot to adapt its gaze behavior to the humans' proxemics around it in order to manage a group interaction. This also gives an insight into how human proxemics can allow the robot to estimate conversational role intentions to understand whether the participant is playing an active, bystander, or overhearer role within the multi-party interaction. This shows the importance of human proxemics as a social signal to adaptation and the potential of non-verbal behavior to display social adaptability.

Second, to inquire into the second question, the thesis discusses the design and implementation of multi-modal behaviors on a robot with respect to the human it is interacting with. The modalities making up the multi-modal behaviors are: proxemics, gaze, and gesture. Various combinations of these multi-modal behaviors are studied as conditions in a data collection involving 105 participants interacting with an autonomous robot. The human behavioral outcomes are observed and analyzed. Consequently, this thesis contributed to the research question by providing a deeper understanding of multi-modal behavior in robotics and the impact of each modality when performed in a multi-modal manner.

Third, two contributions are demonstrated in this thesis while looking into the third question. First, as RL trainings for robots and HRI are costly and challenging specially during the current COVID-19 pandemic, a simulation for HRI is presented and used for the training. Second, a RL model is proposed and evaluated to answer the challenges and limits of current RL for social robotics adaptations. The RL model proposed is used to let the robot adapt its multi-modal behavior by looking into different combinations of modalities. In addition, the RL model puts forward a reward function that considers both the human's multi-modal social signals and their decisions as well as the cost of high-degree multi-modal behaviors of the robot.

As such, this thesis demonstrates the synthesis of socially intelligent multi-modal human-robot interaction through investigating the effect of multi-modal behaviors of the robot on the outcomes of an interaction and proposes an adaptation and personalisation of these behaviors taking into account the human social signals using reinforcement learning.

1.2 Outline of Thesis

The thesis is divided into two parts: Related Work and Contributions.

1.2.1 Part I: Related Work

In the Related Work part, two chapters are discussed as they cover the background and state-of-the-art on the modality behaviors of social robots as well as an in-depth review on the reinforcement learning methods explored for adaption in social robotics respectively. In Chapter 2, background and literature review on the modalities making up robot and human behaviors in human-robot interaction(HRI) are introduced including proxemics, gaze, gestures, emotional expressions, and the impact of behavior on social intelligence and influence. In chapter 3, reinforcement learning and its methods are explained and an in depth literature review of reinforcement learning applications in social robotics for adaptation is presented. The motivation for writing this chapter is the lack of in-depth reviews on the current work and developments of using reinforcement learning for adaptation in social robotics.

1.2.2 Part II: Contributions

In the Contributions, the three main studies of the thesis are discussed and the key findings highlighted.

Robot Gaze Adaptation using Human Proxemics through the use case of group interactions

In the first study of Chapter 4, a pilot study exploring how one human modality can be utilized by the robot to adapt one of its own modalities. Group interactions highlight this importance by having participants of the group playing different roles such as active, bystander, or overhearer. As such, the Chapter introduces a simple model that uses human proxemics to estimate the intended conversational roles in the groups being formed with the robot and consequently have the robot adapt its gaze modality to better coordinate the multi-party interaction. The findings show that such adaptive behavior have positive behavioral outcomes on the interaction by having participants stand closer to the robot regardless of their roles and rated the adaptive robot higher in perceived sociability. This was a first step towards looking at the bidirectional influence in human-robot relationship and more analysis is needed to investigate how other modalities making up multi-modal behaviors of the robot impact the human social signals and the interaction. This chapter was published as a conference paper as [167]: *K. Tatarian, M. Chamoux, A. K. Pandey and M. Chetouani, "Robot Gaze Behavior and Proxemics to Coordinate Conversational Roles in Group Interactions," 2021 30th IEEE*

International Conference on Robot & Human Interactive Communication (RO-MAN), 2021

Investigating the Synthesis & Effects of Multi-Modal Robot Behavior

In the second study of Chapter 5, a multi-modal interaction focusing on the following modalities: proxemics for social navigation, gaze mechanisms (for turn-taking floor-holding, turn-yielding and joint attention), kinesics (for symbolic, deictic, and beat gestures), as well as social dialogue was designed and implemented. The multi-modal behaviors were evaluated through an experiment with 105 participants in a seven minute interaction to analyze the effects on perceived social intelligence through both objective and subjective measurements. The results show various insights of the effect of modalities in a multi-modal interaction onto several behavioral outcomes of the users, including taking physical suggestions, distances maintained during the interaction, wave gestures performed in greeting and closing, back-channeling, and how socially the robot is treated, while having no effect on self-disclosure and subjective liking. These findings allowed for a deeper understanding on the human social signals within a human-robot interaction and how they are impacted by the behavior of the robot. This set an important foundation to build the adaptation and personalization layer still needed to achieve socially intelligent interactions with the robot. This chapter was published as a journal paper as [168]: *Tatarian, K., Stower, R., Rudaz, D. et al. How does Modality Matter? Investigating the Synthesis and Effects of Multi-modal Robot Behavior on Social Intelligence. Int J of Soc Robotics (2021).*

Adaptation of Multi-Modal Behaviors using Reinforcement Learning

Chapter 6 first presents a simulation setup for HRI to enable cost-effective training of reinforcement learning (RL) algorithms in social robotics applications as physical HRI training can be very time consuming and challenging. The simulation design is based on the data collection done in Chapter 5. In addition, the second contribution of Chapter 6 is a proposed generalized adaptation and personalization RL model, where the reward function is designed taking into consideration the human multi-modal social signals, their decisions, and a cost function for high-degree robot multi-modal behaviors. The RL model is investigated through ϵ -MAB, ϵ -UCB, and Q-Learning algorithms trained on the simulation setup. The learning curves of these algorithms as well as the percentage of times the various combinations of multi-modal behaviors of the robot were selected by the RL agents are studied. The findings highlight the importance and priority of proxemics modality in a human-robot interaction. In addition, the findings show that the roles of gaze and gesture modalities

seem to be similar and may be interchangeable. The resulting generalized adaptation RL model has room for personalization for different user profiles and social situations, utilizes the bidirectional influential relationship between the robot and the human user, and demonstrates a promising solution to achieve multi-modal socially intelligent human-robot interaction.

Discussion & Conclusion

Finally, Chapter 7 concludes the thesis by discussing and highlighting the key findings of each study and their limitations. It also sets forth possibilities of future work.

Part I

Related Work

Chapter 2

State-of-the-Art

This Chapter contains:

2.1	Proxemics	11
2.2	Gaze	13
2.3	Gestures	15
2.4	Role Emotional Expressions in Decision Making	16
2.5	Multi-Modal Social Behaviors	17
2.6	Behavior and Social Intelligence and Influence	17

This chapter introduces and highlights the work done in HRI for gaze, kinesics, specifically gestures, proxemics, emotional expressions, and multi-modal behaviors. This chapter also discusses relative work on the effect of behavior on social intelligence and influence.

2.1 Proxemics

In 1959, the term proxemics was defined by anthropologist Edward Hall [67] and he suggested that the zones of proxemic distances are shaped by the culture and *psychophysical* features, which are distance, sociofugal-sociopetal, visual, voice loudness, kinesthetic, olfaction, thermal, and touch [68], [66], [65]. On the other hand, Mehrabian [116] analyzed proxemic behavior and provided distance and orientation metrics between two individuals. In his interest for spatial orientational alignments, Kendon [86] believed that a person's postural shifts, rhythmic organizations, and other actions provide further knowledge that can inform interlocutors of their intentions. In addition, Kendon studied organizational patterns of social encounters and defined the term F-formation, which is when two or more people adjust their spatial and orientation relationship in order to have equal, direct, and exclusive access. Furthermore,

Schegloff [148] also linked proxemics to the intentions of the participants. He used the poses of the lower and upper parts of the body to derive the intentions, whether the involvement of the participant would be dominant or subordinate. Moreover, as defined by Kendon [86], proxemics is affected by the culture of the participants, so it is no surprise that the investigating in the effects of personality and culture on proxemics for human-robot interaction is gaining popularity. After a study on human-robot interactions with multiple participants from three different nationalities, Chinese, American, and Argentinian, has concluded that the cultural backgrounds of the participants have different affects based on the position of the robot [82]. Additionally, in [13], nonverbal behavioral features including proxemics and postures are used to detect extroversion in the personalities of the interlocutors of the robot.

There is a current trend in HRI to not only use proxemics to ameliorate social navigation, but to also use it to initiate interactions in a more natural manner. For instance, Kruse et al.[99] and Dautenhahn et al. [41] used Hall's theory of proxemics to optimize social navigation of the robot while taking into account the human's safety and visibility. Additionally, proxemics can be used to initiate interactions. Shi et al.[151] proposed a model based on proxemics and navigation to initiate a conversation with a human inspired by the study of human-human interactions. The robot with the implemented proxemics model was ranked higher in a subjective evaluation of appropriateness of initiation. An understanding of proxemics grants the robot a finer tool to perceive, predict, and manipulate the environment around the interaction and provides greater naturalness [151]. More study on the initiation and termination of interactions was done in Mead et al. [114] by extracting features to analyze proxemics and recognizing spatiotemporal behaviors.

Moreover, one important application and use case of proxemics has been to understand group formations. Proxemics and participation frameworks of groups have been further studied in human-agent interactions whether the agent is a robot or a virtual agent. For instance, for virtual agents, there is an interest to better understand the set-up of a conversation in a multi-party interaction conveyed using nonverbal signals [103, 131]. In Lee et al. [103, 102], human interactions were studied to model how the different roles of a multi-party conversation affect the behavior of the participants and applied their findings to model the behavior of active participants and bystanders. In addition, in [131], it was shown how the modalities of proxemics and gaze in virtual agents can indicate and signal their conversational roles, also known as footing [60], in groups such as the speakers, addressees, bystanders, and overhearers. Furthermore the need to compute to footing and proxemics, especially F-formation, has expanded beyond doing so for a specific agent. In [14], multi-party interactions were recorded and used to model the participation framework, which is the conversational set-up of an interaction; on the other hand, [111] observed humans to provide a model of the F-formation proposed by Kendon done in [86].

Proxemics is a very rich modality that does not only contain social information such as intentions [116, 65], but it also grants one the ability move and manipulate the physical space around it. Even though navigation is a basic robot skills today, there is still a need to investigate how the synthesis of the robot’s social navigation, influenced by the theories of proxemics, can impact human perceptions of the robot and their behavioral outcomes in an interaction. In addition, there is a necessity to further explore how the human’s proxemics can be used to allow the robot to have a better understanding of the interaction taking place around it. This thesis, first, demonstrates how human proxemics can be used to give insight to the robot on the conversational groups around it and adapt its behavior accordingly. Second, this thesis also investigates how proxemics can be implemented within a multi-modal behavior and the effects it has on an interaction when combined with different modalities. Third, this thesis also presents how the proxemics of the human can be used as a social signal among others for the robot to learn and adapt its multi-modal behaviors, which also includes the proxemics of the robot.

2.2 Gaze

In 1967, Kendon [85] was the first to classify and analyze gaze aversion in human-human interaction, claiming that humans in fact do not spend the majority of their time in a conversation directing gaze straight at another human’s face. He concluded that gaze aversion was done for four primary reasons: turn-taking, turn-yielding, floor-holding, and intimacy regulation (used to regulate the level of shared emotional arousal) [86, 178, 171]. Today, gaze mechanism, including gaze aversion, is still a study of interest for social roboticists. Research in HRI has involved conducting studies to better understand social gaze [125, 11], using gaze to reference an object of conversation by joint attention [93, 10], designing gaze cues to modulate group conversation [125, 126], and regulating turn-taking in conversations [11, 126]. In addition, conversational social gaze constructed of gaze aversions to perform role-signaling, turn-taking, and topic-signaling prompted high indices of likeability towards the robot [125]. Moreover, for robots which lack expressive eyes, head controlled tilts have been designed to convey gaze aversion [11]. The former concluded that while social gaze aversions did not increase the human’s comfort in eliciting more self-disclosure, it did decrease interruption time caused by the user and the robot was perceived as more thoughtful. Additionally, the study analyzed the direction of gaze aversions in human-human interactions with respect to its three primary functions: cognitive, intimacy-regulation, and floor management [11].

An additional important function of gaze is joint attention. It supplies people with a way of interpreting and predicting each other’s actions and focus attention [64]. For instance, speakers tend to use deictic expressions followed by a glance towards the object of reference

[38]. Thus it is no surprise that joint attention attracted the attention of researchers in the HRI field. For instance, it was shown in [22] that users reached objects faster when they could follow the gaze of the robot iCub, who was giving instructions while glancing at referenced object. Similarly, [157] showed that users interacting with a robot that had a gaze with a reference function found it easier to complete a task than with a robot that had random gaze. Joint attention has proven to be functional for social robots to shift the human's attention to the spot at which it is looking [188]. In addition, for collaboration tasks involving object selection, robot gaze shifts assisting its speech were shown to be advantageous for cooperation specially when the human was required to choose the object being referred to by the robot as fast as possible [2, 23]. Furthermore, in [121], during hand over tasks, users started reaching for an item much sooner when a robot consistently gazes at the handover area than when it gazes away from that point. In parallel, when gaze is used as part of a multi-modal behavior, it often has a supportive and enhancing role to other social behaviors, notably speech and gestures [4].

Furthermore, gaze is also used in HRI to allow the robot to adapt to the participation framework in group interactions. In more recent work, it was shown that adapting robot's gaze orientation in group interactions leads to feelings of inclusion by the participants as well as a sense of belonging to the robot's group [176]. In the previous study, the robot would either look at the middle of the group formed or give attention to the participant who was speaking. Moreover, an adaptive robot gaze behavior could shape group interactions with the robot guiding more even participation during collaborative games [59, 127]. For that, the adaptive gaze behavior pattern was formed based on the speech participation of the two participants interacting with the robot and forming the group [59]. While in Multu et al. [127], a social robot used only gaze to signal the roles of the participants whether the role was bystander, addressee, or over-hearer. However, the roles in that study were given to the robot as such there was no autonomous role estimate done by it.

While there has been and still is on going research on the effects and functions of gaze mechanisms for social robots in HRI, most of them are done independent of other modalities that can be present on the robot such as gestures and proxemics. Thus there is a need to investigate the influence of gaze functions when the robot is performing in multi-modality to better understand the extend of the studied effects and how can it be adapted based on the social signals of the human the robot is interact with. This thesis studies the extent of which gaze can be used to have the robot adapt to changes of human signals, specially in the case of conversational roles in groups. The thesis also designs and implements four gaze mechanisms for a social robot for the functions of turn-taking, turn-yielding, floor-holding, and joint attention and examines the effects of these mechanisms when combined with different modalities making up the behavior of the robot.

2.3 Gestures

Gestures for humans have been categorized and defined primarily based on their role in communication and their functions as follows [113, 97, 6, 48]:

- *Iconic gestures* for describing physical objects and events mentioned in a conversation; e.g., forming a small circle with the hand to refer to a small ball.
- *Metaphoric gestures* for depicting abstract concepts being referred to; e.g., fast back-forth hand movement to indicate ‘ongoing’ work.
- *Deictic gestures* for indicating objects in the physical space where the conversation is taking place; e.g., point at a road close by.
- *Emblem gestures* or symbolic gestures for expressing language-like features with agreed upon culturally specific properties; e.g., the V hand gesture with the index and middle fingers to indicate a peace sign.
- *Beat gestures* for emphasizing significant points or certain words in the speech using rhythmic movements of hands and arms e.g., hand gesture to indicate the introduction of a new topic

Gestures have also been studied and implemented on robots aiming to improve human-robot interaction. While deictic, beat, iconic, and metaphoric gestures were all found to boost the robot’s performance as a narrator in a narrative scenario, deictic gesture significantly ameliorated the user’s recall of information on the story [79]. Additionally, the robot which performed correctly timed nods in a conversation and proper gaze and gesture sharing behaviors was ranked more highly than a robot who did not have such behaviors [83].

Moreover, gestures play a role in portraying emotional expressions. For instance, submissiveness can be expressed by an open hand shape; on the other hand, dominance can be portrayed in a pointing hand shape [92]. Similarly in social robotics, modulating the robot’s body movement by varying its head tilts and body expansiveness influenced perceived dominance [132].

Another important aspect of social interaction is *alignment*, which refers to the convergence of linguistic behavior and/or similarity in mental representation [134, 24]. Alignment is an ubiquitous feature used to measure to which extent interactions shape behavior and their success at communicating shared understanding [72]. For instance, it was shown that alignment, through mimicry of postures, mannerisms, and facial expressions in dyadic interactions (chameleon effect [35]), increased the rapport between the participants, the pro-social behavior even beyond the interaction and smoothness of interaction [172]. Moreover, a study found that users, when retelling a story to a third participant, were more likely to

demonstrate the same iconic gestures they witnessed the first time [120]. Alignment equally plays an important role in human-computer interaction in enhancing communicative success [24]. In robotics, people have been found to nod more when interacting with a robot that nods along in response to that in comparison to a robot who does not mirror their nodding [152]. In addition, a computational method for evaluating and modeling of interpersonal synchrony in behaviors during interactions offered a perspective for building social interfaces for robots and embodied conversational agents [46]. Furthermore, motor resonance, which is the activation of the observer's motor control system during action perception, was used to not only produce more natural interactions for robots with humans but also as an evaluation method to determine quantitatively how the robot is perceived by the human [149].

2.4 Role Emotional Expressions in Decision Making

In the recent years, there has been growing interest in research on the interpersonal effects of emotion, specially for autonomous socially intelligent machines. As it has been shown that such machines can increase cooperation with human users by simulating emotional expressions as well as perceiving and comprehending human users' emotions [63]; thus influence their decision making. While there are few research studying robotic emotional expressions [44, 28, 91, 84, 71, 186], it is first important to highlight some of the fundamental theories investigating how emotional expressions may influence behavior. For instance, according to the appraisal theory [122, 158], a human's emotions are a result of their evaluation and assessment of a certain situation, as such it may lead to different reactions for different people and situations. On the other hand, the reaction of the observer takes place through reverse appraisal [42]. Moreover, according to Emotions As Social Information (EASI) of Van Kleef [94], emotional expressions are used to modulate social interactions and shape behaviors, specially for establishing cooperation between individuals [173].

In Human-Computer Interactions (HCI) through virtual agents and even anthropomorphic robots, emotional expressions in cooperative and competitive relationships have been studied to evaluate the impact on the decision making of the human. For instance, the emotional expression of joy allows the actor to inform the observe that it favors the current situation and would like the observe to continue their current behavior thus leading to a possible cooperative relationship [31, 124, 138]. On the other hand, display of the emotional expression of anger has shown mixed results; while for some studies it was used by the actor to show the observer that they are not in favor of the current situation and the observer's current behavior [52, 137, 155], some studies have shown that anger may not be the optimal strategy to get more compromises in negotiation [39, 154].

In Wang et al. [182], it was shown that the roles of the actors and observers (follower vs. leader) impact how the emotional expressions are received. It also highlighted that persuasion was more probable when the person knew it was appropriate to exhibit emotional expressions. In addition, in Melo et al. [45, 43], a prisoner’s dilemma study demonstrated that the emotional expressions of a virtual agent strengthen cooperation. Similarly, the study was replicated recently with a social embodied robot (a NAO robot) [165]. The emotional expressions of the robot (joy, anger, shame, and sadness) were created using body motion, speech, and colors of the eyes’ LEDs. The results showed that when the robot behaved cooperatively and it displayed anger when participant betrayed it and joy when they cooperated with it, the human was more cooperative. Moreover, in Broekens et al. [28, 118] presented well-founded indications that show the potential advantage of having affective communication with the human users in the reinforcement learning loop.

2.5 Multi-Modal Social Behaviors

Multi-modal behaviors have been studied by various fields beyond HRI including neuroscience and psychology. For instance, it was found that humans rely on several modalities, such as voice, facial expressions, and postures, at a time to perceive emotions [32, 189]. Moreover, in neuroscience, it was shown that the human brain combines information from different sensory inputs in order to strengthen perception and direct behavior [51, 50]. On the other hand, in HRI, several studies have tried to study multi-modal behaviors by combining two or more modality behaviors to assess an interaction. Most commonly, gaze behaviors are combined with gestures [69], proxemics [184, 54] or verbal behaviors [25]. Work combining other modalities [142, 133], or comparing more than two or three modalities [20, 87] are rare. However, the manipulations in these cases typically combine some modalities and evaluate overall system performance, as opposed to investigating the effect of specific modalities. As such, there is a clear need to develop a more comprehensive perspective on how different combinations of modalities (gaze, gesture, proxemics, and verbal content) contribute to overall perceptions of social intelligence during the course of an interaction.

2.6 Behavior and Social Intelligence and Influence

Social influence has been broadly defined to include changes in an individual’s cognition, attitudes, physiological responses, and behaviors that arise due to the presence of another person and the attributions made about that person [7, 70]. Social influence can be manifested in several ways, such as social norms, conformity, which includes shifting one’s attitude or behavior to match another’s, and compliance, which is defined as accepting a request or

desired behavior [37]. As such social influence is closely associated with emotion. For instance, it has been shown promotion and prevention of a certain behavior is attributed to an emotional response, which is linked to goal pursuit and defined as a combination of the dimensions of valence and arousal [75, 27]. Thus, emotional behavior plays a major role in the decision-making of everyday life decisions, especially for promoting ethical choices in purchasing decisions, as shown in [80]. For instance, more recent work in social robotics has been focused on designing emotional expressions and mood gestures for social robots and investigating how they are perceived [132, 186] as well as how these emotional expressions influence decision making of the user [129]. In addition, prior work on robot influence has repeatedly found multi-modal behaviors during HRI to consistently outperform unimodal ones, especially, when influencing people's cognitive frames, emotional responses, and task performance [145]. One recent study on this subject was conducted by [146]. They studied how a robot's persuasive behavior influences people's decision making. The results showed that a strategies based on emotion had significantly higher persuasive influence compared to both the logic based strategy and control conditions. Moreover, in a human-agent interaction, it was shown that the three factors that impact the virtual agent's influence and persuasion using emotional expressions are: the role the agent and person are playing (leader vs. follower), the appropriateness of displaying an emotional expression with respect to the context, and the verbal phrasing and dialogue [182].

Modalities making up the multi-modal social behavior have also played a role in shaping the robot's social presence. For instance, socially aware proxemics models proved to have a positive influence on social presence [54]. However, some studies investigating the effect of social dialogue found no impact for social-speaking versus neutral-speaking robots [106], nor for subjective speaking versus fact-based speaking robots [174] on the robot's social presence. How other modalities may influence the robot's social presence is yet to be investigated. In addition to the modalities of nonverbal behaviors, dialogue plays an influential role in forming impressions and manipulating social outcomes of the interaction. In human-human interactions, one study has shown that starting a conversation by asking people how they were feeling that day increases the likelihood of their compliance to a request for both charity donations and/or commercial purchases [47]. Moreover, in human-agent interactions, it was shown that having an agent start with a small request increased the chances of having the participant accept a bigger request shortly after [47]. Furthermore, verbal phrases influence social interactions with agents, e.g., separating emotional expressions targeted at an attitude versus at a person such as "your opinion" versus "you should" [182]. In HRI, dialogue similarly has an impact on the interactions including facilitating collaborations, managing errors, and personalizing conversations. For instance, it is used to exchange information and assist in human robot collaboration to achieve common goals [56]. Furthermore, social dialogue was

shown to help robots recover from prior errors and gain future influence [110]. Additionally, service robots with personalized dialogues reinforced participants' rapport, cooperation, and engagement [104].

Chapter 3

Review on Reinforcement Learning in Social Robotics for Adaptation

This chapter contains:

3.1	Reinforcement Learning	20
3.1.1	Multi-Armed Bandit	22
3.1.2	Q-Learning	22
3.1.3	Background and Definitions	23
3.2	Reinforcement Learning In Social Robotics	24
3.2.1	RL for Robot Adaptation and Selecting Appropriate Behaviors	25
3.2.2	Exploring Reward Functions for RL in HRI	31

3.1 Reinforcement Learning

In RL, an agent learns the optimal way of interacting with its environment by evaluating its actions to maximize its rewards [161]. The standard reinforcement learning framework is shown in Figure 3.1.

The models and algorithms of RL and their taxonomies are summarized in Figure 3.2. One class of RL is called **multi-armed bandits** (MAB), or **k-armed bandits**, and in which the agent simultaneously tries to gain new knowledge by *exploration* and optimizes its decisions based on existing knowledge by *exploitation*. By balancing *exploration-exploitation*, the agent must find the best sequence of actions yielding the highest rewards and exploit them [177]. Furthermore, another class of RL is based on Markov Decision Process (MDP) and is divided

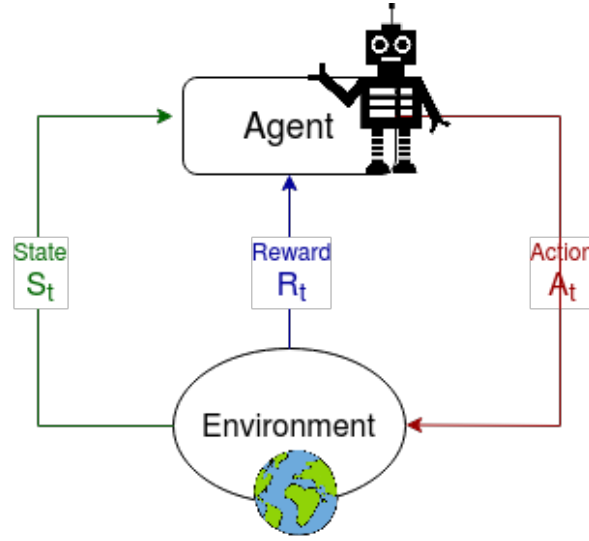


Fig. 3.1 Standard Reinforcement Learning framework (inspired by [161])

into sub-classes **Model-Free** and **Model-Based**, in which the learning algorithms have to explicitly reference the model. On the other hand, **Model-Free** RL can be *Value-Based* or *Policy-Based*, which can be Gradient Free or Gradient Based. This study focuses on MAB's as well as *Value-Based* RL methods, which are further explained below:

Value-Based Reinforcement Learning:

1. **On-Policy:** where the agent learns the value of the policy used to make decisions (target policy and behavior policy are the same) and learns policy π from experience sampled from π (learning on the job). An example of this is **SARSA:** (state-action-reward-state-action algorithm) where the agent interacts with the environment and selects an action based on the current policy.
2. **Off-Policy:** where the agent learns about policy π from experience sampled from μ and uses the experience from old policies or other agents to improve the policy (learning from someone else's behavior). An example of this is **Q-learning:** using discounted reward and applying iteratively the Bellman equation. The target value is not dependent on the policy being used but only on the state-action function. Furthermore, **DQN:** is Q-learning with Convolutional Neural Network (CNN), the agent utilizes an automatic feature extractor through deep CNNs to approximate the action-value function of the Q-learning method.

For the purpose of this study, two reinforcement learning algorithms, Multi-Armed Bandit and Q-Learning, were used to investigate the adaptive multi-modal behavior system.

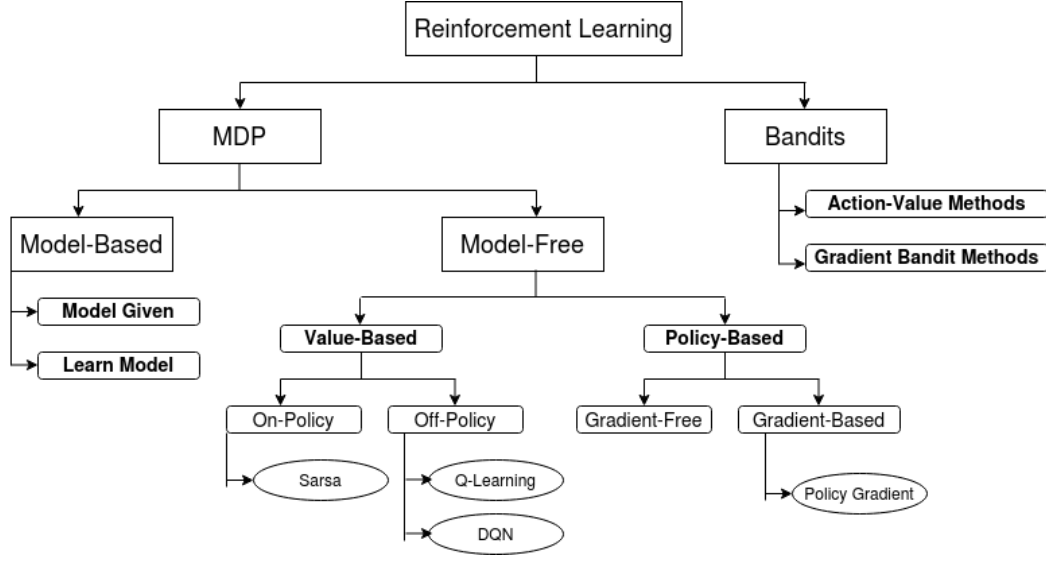


Fig. 3.2 Map of Reinforcement Learning algorithms (inspired by [191])

3.1.1 Multi-Armed Bandit

The Multi-Armed Bandit (MAB) problem, also known as the k -Armed Bandit, is a non-associative, evaluative feedback problem. Each of the k actions has an *expected* or *mean* reward given to that action when it is selected. Then the value of an arbitrary action (a) is described in Equation 3.1.

$$q_*(a) \doteq \mathbb{E}[R_t | A_t = a] \quad \forall a \in \{1, \dots, k\} \quad (3.1)$$

where A_t is the action at time step (t) and R_t is the corresponding reward.

3.1.2 Q-Learning

Q-learning is an off policy value-based RL algorithm and it seeks to find the best action to take given the current state. In Q-Learning, the agent learns from actions that are outside the current policy, like taking random actions, and for this reason it is considered as off policy. More specifically, in Q-Learning the agent attempts to learn a policy that maximizes the total reward. When the agent interacts with the environment, it updates the value the state-action pairs $Q(s, a)$ in the Q-table using Equation 3.2. Initialized at random variables, as the Q-table gets updated, it serves the agent as a reference to select the best action based on the q-values.

$$Q(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q(s_t, a_t) + \alpha \cdot (r_t + \gamma \cdot \max_a Q(s_{t+1}, a)) \quad (3.2)$$

where $Q(s_t, a_t)$ is the value of the state-action pair, α is the learning rate, γ is the discount factor, r_t is the current reward, and $Q(s_{t+1}, a)$ is the value of the action-state pair in the next state (s_{t+1}).

3.1.3 Background and Definitions

It is important to go over and highlight some crucial concepts and parameters that will be used and manipulated:

- *Learning rate* (α): or step size regulates to which extent newly obtained information overrides old ones. A factor of 0 makes the agent exclusively *exploit* prior knowledge without any learning. A factor of 1 makes the agent ignore prior knowledge to *explore* new possibilities.

$$Q_t(a) \leftarrow Q_t(a) + \alpha(R_t - Q_t(a)) \quad (3.3)$$

- *Discount Factor* (γ): regulates the importance of future rewards. A factor of 0 makes the agent short-sighted, *myopic*, by only taking into account current rewards. On the hand, a factor of 1 makes an agent aim for long-term high rewards. Thus, choosing a high gamma value would lead to an agent that considers the information to be obtained from the next state more important than the current one. For adaptation, an ideal discount factor would high so that the robot adds more value to the future rewards so that its decisions optimally lead to adapted and personalized behavior. In this study, the discount factor was set to $\gamma = 0.95$.
- *ϵ -greedy Action Selection*: An action k is selected at random (with uniform probability) for a proportion ϵ and an action with the highest estimated values is selected for a proportion $1 - \epsilon$ of the trials:

$$A_t = \arg \max_a Q_t(a) \quad (3.4)$$

- *Upper-Confidence-Bound with ϵ -greedy Action Selection*: where uncertainty is used to drive exploration. In this algorithm, the agent explores more at the beginning to reduce uncertainty then exploration decreases. This method is not ideal for nonstationary problems and large state spaces. Similarly to above, an action k is selected at random (with uniform probability) for a proportion ϵ and for $1 - \epsilon$ of the trials:

$$A_t \doteq \arg \max_a [Q_t(a) + c \sqrt{\frac{\ln t}{N_t(a)}}] \quad (3.5)$$

where $Q_t(a)$ is the highest estimated value, c controls the amount of exploration, $N_t(a)$ is the number of times the action is taken, and $c\sqrt{\frac{\ln t}{N_t(a)}}$ is the upper-confidence bound exploration term.

3.2 Reinforcement Learning In Social Robotics

In RL, an agent interacts with its environment by trial-and-error in order to achieve an optimal behavior. Thus making RL a framework for decision-making problems. Similarly to RL, interaction is also an integral part in social robotics. For this reason, RL is an appropriate method to tackle real-world interactions and problems for physically embodied social robots and human-robot interactions. In an article surveying reinforcement learning (RL) approaches in social robotics, RL approaches in HRI were categorized based on the used method and the design of the reward mechanisms [5]. The three categorizes were presented as follows:

1. Interactive RL: The feedback received by the human user is done through evaluative feedback, corrective feedback, or guidance. This feedback can be given in two ways:
 - *Explicit Feedback*: After observing the agent's actions, the human teacher gives feedback to the agent either by direct manipulations, such as using the robot's touch sensors, or through an artificial interface or GUI.
 - *Implicit Feedback*: The learning agent is trained by the human teacher indirectly through a natural and spontaneous interaction. One of the biggest challenges is that can be noisy.
2. Intrinsically Motivated Methods: The social robot tries to maintain and reach an optimal internal state by evaluating internal and external factors.
3. Task Performance Driven Methods: *Human task performance*, *Robot task performance*, and *human and robot task performance*

Furthermore, in *Interactive RL*, the feedback provided by the human teacher is integrated to form the action policy, by being involved in the action selection mechanism, or to model the reward function [5]. There are three categorizes of evaluative feedback: first is *interactive shaping*, in which the feedback given by the human can be interpreted as a numeric reward; second is *learning from categorical feedback*; and third is *learning from policy feedback*. Thus by narrowing the action choices of the robot, the human instruction provided can shrink and in turn making the training process faster to converge to the optimal path [5]. This study focuses on Interactive RL through implicit feedback using multi-modal social signals from

the human. Social signals may change continuously and as such cannot be discretised as explicit feedback. Before further diving into the proposed model, relative work on using RL in social robotics needs to be discussed.

3.2.1 RL for Robot Adaptation and Selecting Appropriate Behaviors

Some studies and literature in HRI have tried to implement and evaluate adaptation using table-based algorithms and thresholding but faced limitations. For instance, Tanevska et al. [166] introduced a socially adaptable framework, where the robot's adaptation is guided by an internal social motivation system. The robot takes into account input from the behavior of the person it is interacting with by checking visual and tactile stimuli and then accordingly adapts its behavior to engage and disengage from the interaction to reach optimal internal comfort for itself (the robot). The comfort of the robot depended on the intensity of the interaction with the human. For instance, if it was left along for a prolonged period of time, the comfort level would decrease to critical. The implemented model is shown in Table 3.1.

Table 3.1 Summary of Adaptation Model using Thresholding in Tanveska et al.(2020)

Objective	Internal social motivation system for robot to maintain optimal comfort level
Method	Thresholding with parameters of the decay and growth rates calculating the comfort value in relation to the previous comfort value level (if there is visual ($F(t)$)/tactile stimuli ($T(t)$): $C(t) = (F(t) + T(t) + C(t - 1) * \tau) / (\tau + 0.1)$ else: $C(t) = \beta * C(t - 1)$. Start with comfort level at 50% of maximum level.
Model	During interaction, the behavior of iCub was a result of a state machine with 3 main states: idle, interact, and suspend; and 4 actions: moving its arms, looking, straighten up action, and lean down actions.
Feedback & Input	Visual and tactile stimuli from human user

The system was evaluated with the iCub robot in a physical interaction. However, results showed that not all scenarios were equally explored as the critical level was much more dominant than the saturation level (high intensity interactions with the robot). In addition, while the *Thresholding* approach for adaptation is well suited for behaviors that are linked to specific signals and allow for good reactive actions to that continuous flow of data, it does not allow for complex reactive and dynamic behaviors for adaptation nor

is it beneficial for long-term understanding of the user [100]. Thus there is a need for a method that would allow for such dynamic and reactive behaviors while still taking into account previous experiences with the user and exploring as well as exploiting; a possible solution for this is RL. The rest of this subsection will discuss relevant work of RL in social robotics for adaptation and selection of appropriate behavior in human-robot interactions.

Bagheri et al. [17] presented a framework to enable cognitive empathy in social robots. The robot first uses facial emotion recognition to perceive affective states of human users, the robot is later provided with a RL model so that it learns the most appropriate empathetic behaviors, which in this study were empathetic utterances, of different states based on the emotional feedback response of users. The summary of the model, states, actions, and reward is presented in Table 3.2.

Table 3.2 Summary of RL Model in Bagheri et al.(2020)

Objective	Robot learns appropriate empathetic utterances
Method	RL Q-learning, with context bandit, to learn optimal action-selection policy ($\epsilon = 0.1$ for ϵ -greedy action selection, learning rate $\alpha = 0.1$, and no discount factor)
States	12 States = 4 Emotions (<i>Anger, Sadness, Happiness, Surprise</i>) X 3 Personality Types (<i>Introvert, Ambivert, and Extrovert</i> based on <i>extroversionscore</i> by <i>personality questionnaire</i> by [62])
Actions	4 Possible Empathetic Utterances pre-defined (<i>Mimical, Motivational, Distractical, Alleviational</i>)
Reward	$R = +1$ if state changes to neutral/happiness/surprise & $R = -1$ if state changes to anger/sadness

At the beginning, the Q-table was initialized all with zeros and thus the algorithm selected random actions. The updating of the Q-table relied on two possibilities. First, when found in a "undesirable" or negative state (i.e anger/sadness), the Q-value of that selected action decreased. Second, if the state is "desirable" (i.e neutral/happiness/surprise), the Q-value of that selected action increased. Evaluation study took place as a collaborative game with 28 participants, which were classified based on the three personality types and the emotional recognition was done using Wizard-Of-Oz technique. The results showed that the negative emotions were not detected often and thus the employed RL did not converge to the empathetic utterances that could have impacted the affective state of participants. The evaluation

was also done through questionnaires and showed higher ratings in Social Presence, Perceived enjoyment, and engagement [17]. While this study relied only on emotional recognition for the learning of the robot and created a large state-space, the employed RL did not converge as some states were not visited. A possible solution for this is looking at other social signals as well.

Hussein et al. [135] proposed a Multimodal Deep Q-Network (MDQN) for the robot to learn human-like interaction skills through a trial and error method and using end-to-end reinforcement learning. A Pepper robot was placed in an open-space, where people can freely pass by, for 14 days to learn appropriate timing to handshake people. The input included *grayscale frames* from top 2D camera of Pepper, *depth frames* from 3D camera, and *touch sensor* from an external glove added on the robot's right arm. The MDQN consisted of two streams of Convolutional Neural Networks (NN) for action-value estimation. The dual stream converts process the depth and grayscale images independently. The Algorithm divided into two phases to avoid delay:

- **Data Generation Phase:** (during interaction) generations a memory that keeps the most recent experiences (s_t, a_t, r_t, s_{t+1}) to be used in training phase for updating the network parameters.
- **Training Phase:** (during rest) from each experience replay, a mini buffer is randomly sampled from memory and on which the model is trained. the network parameters are updated iteratively in the direction of the bellman targets.

Table 3.3 Summary of RL Model in Hussein et al.(2017)

Objective	Robot learns appropriate time to handshake
Method	Multimodal Deep Q-Network (MDQN) using end-to-end reinforcement learning
Actions	4 actions: waiting, looking towards human, waving, and handshaking.
Reward	$R = +1$ if handshake attempt was successful, $R = -0.1$ if the handshake attempt failed, and $R = 0$ for all other actions

The summary of the model, actions, and reward is presented in Table 3.3. First to evaluate Model Performance, a separate test data-set was collected and used to evaluate decision-making of the robot. Three volunteers judged the agent's decision on whether the action taken was appropriate or not. If the agent's decision was considered wrong by the majority, then the volunteers were asked to agree on the most appropriate action for that particular

scenario. Second to validate use of two streams to improve social cognitive ability, accuracy corresponding to how often the predictions by the Q-networks were correct was checked. The networks compared were y-channel Q-network, depth-channel Q-network, MDQN on the test data-set and the accuracy, true positive rate, false positive rate, and misclassification rate were checked. MDQN had the highest accuracy with 95.3% . In addition, the learned factors that formed the bases for intention inference were activity in progress, walking trajectory, and head orientation. Furthermore, for designing the reward function, five different networks with five different reward functions, where for unsuccessful handshake the rewards were 0,-0.1,-0.2,-0.5 were trained. The plot of the true positive rate of each model on test dataset versus corresponding penalty given on unsuccessful handshake showed that the -0.1 penalty achieved maximum accuracy on the test dataset [135].

Using a different modal of RL, Gao, Sibirtseva, et al. [58] presented a meta-learning policy gradient method as an adaptation solution for HRI and to investigate its role in modelling trust. The interactions were modelled as *adversarial multi-armed bandit* (MAB) problems. In addition, to address the sample inefficiency problem of policy gradient method using a meta-learning algorithm, *model-agnostic meta-learning* (MAML) [53] was used. The experimental set-up was an escape the room scenario in a virtual-reality set and the players were locked in the room with the robot and needed to solve a series of puzzles in order to escape it under a time constraint. An important assumption made is that the interactive space C is made of conation, cognition, and affection $C \equiv \{T, S, A\}$ (inspired by Hilgard's modern behavior psychology [76]). The robot's behaviors and actions consisted of only verbal replies. Additionally, the Meta Learning technique was made of $M = \{\zeta_p, \zeta_r\}$, where ζ_p is the pre-training method with meta-policy $\pi_{meta} = \zeta_p(\epsilon_x, \pi)$ and ζ_r is the refinement method. The interactive space was further designed as follows:

1. The human feedback of each action of MAB, based on the human's position in the maze, was modelled as a Gaussian distribution $r^c \sim N(\mu^c, (\sigma^c)^2)$ for all auxiliary environments $(\epsilon_x^{c \in \{T, S, A\}})$.
2. ζ_p^c to train initial random π^c in order to get a meta policy π_{meta}^c
3. Human experiments were conducted and studied with different subjective measures along with interaction

For implementation, MAML was implemented in PyTorch, trust region policy optimization (TRPO) as optimization algorithm for all policies, and MAB environment implemented with OpenAI Gym. Table 3.4 summarizes the work and the study of [58]. Results from the study show that the algorithm presented in this paper had a higher trust and faster adaptation in

comparison to an algorithm based on Exp3 (Exponential-weight algorithm for Exploration and Exploitation; a type of Bandits Algorithm) .

However, the findings are limiting as only the verbal reply of the robot was manipulated and its embodiment was not used. Moreover, the adaptation proposed is task-dependent as the only feedback is the position of the user within a maze they are trying to escape.

Table 3.4 Summary of Meta-RL adaptation for trust in Gao, Sibirtseva, et al.(2019)

Objective	Adaptation in HRI and Investigation of bidirectional trust
Method	Policy gradient based solution for MAB problem with meta-learning
Actions	12 actions in total: 4 actions, which were verbal replies, for the 3 interaction instances: conation, affection, and cognition. Each instance was triggered by participant's position via the Hololense → robot chooses 1 of 4 actions(replies)/instance according to a probability distribution provided by algorithm.

Moreover, in some RL work for social robots, manual input from humans were used to train the robot. However, such studies would be well-suited to answer the adaptation problem as the training for behaviors was long and computationally expensive and required the human to give explicit feedback. For instance, Knox et al. [96] presented and used TAMER(Training an Agent Manually via Evaluative Reinforcement) to enable the training of interactive navigation behaviors on the Mobile-Dexterous-Social (MDS) robot "Nexi"[95]. The summary of the model, states actions, and reward is presented in Table 3.5.

TAMER agent is made of:

1. **credit assigner**, which takes in action a , delayed reward h by human, state s , and output samples (s, a, \hat{h}) . The credit assigner addresses the small delay in providing feedback by spreading each human reward signal among multiple recent state-action pairs, contributing to the label of each pair's resultant sample for learning $R\hat{h}_H$. The probability density function f_{delay} used here was Uniform distribution.
2. **supervised learner**, which inputs the (s, a, \hat{h}) , outputs the reward model $\hat{R}_H : S \times A \rightarrow \mathbb{R}$, and uses a regression algorithm chosen by agent designer. In this case study, the \hat{R}_H was modeled using k-nearest neighbors with a separate sub-model per action.
3. **action selector**, which inputs the state s from environment and \hat{R}_H and outputs the action a . The TAMER agent acts greedily by choosing the action $a = \operatorname{argmax}_a [\hat{R}_H(s, a)]$.

(this is equivalent to performing RL with a discount factor of 0, where reward acquired from future actions is not considered in action selection - *myopic*).

From the evaluation of the system, nearly all unsuccessful sessions failed due to a lack of transparency, which was related to the mismatches between the state-action pair currently occurring and what the trainer believes to be occurring due to the visibility range by the infrared camera. The *go to* behavior was taught successfully early on with active training time of 27.3 minutes; however, the transparency issues disabled any further success. For this reason, for the remaining four behaviors, an out-of-range alarm was added. The 4 behaviors were taught successfully in consecutive training sessions and a shorter amount of time (9.5 min, 5.9 min, 4.7 min, and 7.3 min).

Table 3.5 Summary of RL Model "TAMER" by Knox et al. (2013)

Objective	<p>Use TAMER for the robot to learn 5 interactive navigational behaviors from the feedback signals provided by the human trainer:</p> <ul style="list-style-type: none"> • <i>Go To</i> • <i>Keep Conversational Distance</i> • <i>Look away</i> • <i>Toy Tantrum</i> • and <i>Magnetic Control</i>
Method	TAMER agent consisting of a credit assigner, supervised learner, and an action selector
States	2 state features: distance between robot and object and orientation angle θ between robot and object
Actions	4 actions: turn left, turn right, move forward, and stay still.
Reward	Through remote buttons provided by the human giving +1 and -1 rewards

Social interactions, specially even human-robot interactions, are dynamic, multi-modal, and complex as the robot itself is an embodied agent capable of multi-modal behavior. Previous work have mostly focused at each time one part of the interaction, either by making the adaptation task-dependent or by generating only one type of robotic behavior or even by only looking at one feedback from the human. Humans interact with the environment by using social signals and in many dimensional ways. One way of capturing that would require having very large state-spaces, which would in turn require a long learning time as well as heavy computational efforts. On the other hand, another way to try and solve the problem is to explore the reward function of the RL model.

3.2.2 Exploring Reward Functions for RL in HRI

RL-based approaches, in which the reward functions capturing social signals in real-time during the interactions, were studied with an objective to have the robot quickly learn about and dynamically adapt to individual human preferences.

In Ritschel et al. [139], a reward signal, which was calculated straight from the human's engagement, was used to try and optimize the user's engagement as well as adapt the robot's personality. The robot's personality was demonstrated through its linguistic style by producing utterances with different degrees of extraversion using Natural Language Generation (NLG). Moreover, the **user engagement score** E_t was calculated based on the multi-modal social signals captured by a Microsoft Kinect 2 sensor and analyzed by the Social Signals Interpretation (SSI) framework [180], which processes and interprets the data in real-time. Once the head, gesture, and posture features were extract, the user's engagement E_t was estimated based on the Dynamic Bayesian Network (BN). The BN is a directed, acyclic graph with nodes representing variables and edges describing conditional probabilities. The BN used in their system was modeled with GeNIe. The probabilities of the variables in the network were learned based on the NoXi corpus, which includes interactions of experts and novices about a certain topic, including Audio, Video, and Kinect 2 depth streams. The measurements used during the interaction were:

- **User Engagement** at time t (E_t):

$E_t > 0$: user engaged

$E_t < 0$: user not happy with the interaction

- **Change in User Engagement**: ΔE_t ($\Delta E_t = E_t - E_{t-1}$)

$\Delta E_t > 0$: level of extraversion close to user's preference

- **Robot's Extraversion(X)**: as a discretized value in the integer interval $[-2; +2]$ translated as *[very introvert, introvert, neutral, extravert, very extravert]*. X then influenced the NLG parameters.

Table 3.6 Summary of Proposed Simulated RL Model by Ritschel et al. (2018) [17]

Objective	Dynamically adapt to individual human preferences in a story-telling setup
Method	RL Q-learning with ϵ -greedy ($\epsilon = 0.2$, exploration high enough to handle noise; $\alpha = 0.5$ learning rate low enough to not eliminate all previous knowledge in case of noise)
States	2D state-space: [estimated user engagement, robot's current extraversion] ($[E_t, X]$) (as integer values ranging in $[-2; +2]$)
Actions	<p>3 actions to control the robot's personality:</p> <ul style="list-style-type: none"> • Increase X by 1 • Decrease X by 1 • No change <p>The actions were chosen to prevent the robot from changing X too fast and to learn faster with less actions</p>
Reward	<p>based on ΔE_t and task independent:</p> <ul style="list-style-type: none"> • $\Delta E_t > 0$: Positive Reward • $\Delta E_t < 0$: Negative Reward (Punishment) • $\Delta E_t = 0$: reward $R = +0.5$

The summary of the proposed model, states actions, and reward is presented in Table 3.6. At each description (step), when the robot stops speaking, a new user engagement E_t was used to calculate the reward based on $\Delta E_t = E_t - E_{t-1}$. The simulation consisted of 30 learning steps and started with neutral extraversion $X = 0$ and an empty Q-Learning table. To evaluate how much would be needed to adapt to the preferences, changes for the user preferences were introduced at time steps 15 and 26 to random values all in simulation.

Moreover, two separate reinforcement models were investigated to create an adaptive multi-modal joke generation/humor of a social robot [183] [140]. Both RL approaches used a linear function approximator with normally distributed initialized weights and then computed the action value and updated it iteratively using gradient descent to find the optimal weight vector; an algorithm based on Sutton et al. [162]. To find an approximation of the optimal action value function $Q^*(s, a)$, which is approximately the weighted function $Q(s, a, \omega)$, the linear function approximation allowed to calculate the weight vector ω for every action $a \in A$ and use the equation 3.6, where $\phi(s)$ was the representation of the current state 3.6:

$$Q(s, a, \omega) := \phi(s) \circ \omega, \forall s \in S, \forall a \in A \quad (3.6)$$

In both approaches [183] and [140], a Social-Signal-Interpretation Framework (SSI) [180] was used to capture and analyze the necessary audio-visual signals from the human interacting with the robot. The audio and the visual signals were used to calculate the estimated probability of the human's vocal laughter and the human's smile respectively Table 3.7.

While Ritschel et al. [140] ran a simulation of the user's preference for noise probabilities at 0% (baseline), 5 %, 10 %, and 30 %, stability at $R = 0.5$ was reached with no noise and the overall performance decreased with the added noise as randomness in reward but coping with the noise was maintained. On the other hand, Weber et al. [183] ran a within-subjects user study with 24 participants with two conditions a learning robot versus a random non-learning robot. While there was no difference between the detected amusement level percentage reward between the learner and nonlearner across all sessions, there was a carryover effect when in the first session the robot was a learner [183].

Table 3.7 Summary of Weber et al. (2018) and Ritschel et al. (2020) for their approach in using RL to have an adaptive multimodal joke generator for a social robot

Studies	Weber et al. (2018)	Ritschel et al. (2020)
Objective	Keep the user's vocal laughs and facial smiles as high as possible	Use human's laughter and smiles to optimize the use of the robot's multi-modality for the next joke
Method	<p>RL with linear function approximator and gradient descent was used to find optimal weight vector of Equation 3.6 using mean-squared error (MSE) as objective function, the derived of which is the follow Equation 3.7:</p> $\Delta\omega = \alpha \sum \mathcal{P}_s(Q * (s, a) - Q(s, a, \omega))\phi(s) \quad (3.7)$ <p>based on Sutton et al. [162]</p>	<p>RL with linear function approximator using Equation 3.6. In order to calculate the optimal weight vector ω, the reward \mathcal{R}_t is used to update the weight vector ω_t until convergence to the optimal one using the following Equation 3.8:</p> $\Delta\omega_t = \alpha(\mathcal{R}_t + \gamma \max Q(s_{t+1}, a_{t+1}, \omega_t) - Q(s_t, a_t, \omega_t))\phi(s_t) \quad (3.8)$
States	<p>2D nondiscretized vector ϕ:</p> $\phi = \begin{pmatrix} \mathcal{P}(\text{usercurrentfacialsmile}) \\ \mathcal{P}(\text{usercurrentvocallaughter}) \end{pmatrix}$ <p>where $\mathcal{P}(\text{usercurrentfacialsmile})$ and $\mathcal{P}(\text{usercurrentvocallaughter})$ are the probabilities of the estimated facial smiles and vocal laughter of user detected respectively</p>	<p>state s_t defined as a 4-tuple $(\text{pitch}, \text{speech rate}, \text{volume}, \text{break}) \in \mathcal{S}$, which is then converted into the vector $\phi(s_t)$, which is divided in four sections:</p> <ol style="list-style-type: none"> 1. <i>Pitch</i>: low, medium, high 2. <i>Speech rate</i>: slow, medium, fast 3. <i>Volume</i>: soft, medium, loud 4. <i>Break (pause)</i>: short, long
Actions	The action was selected by actions made up of 23 sounds, 19 grimaces, 3 types of jokes (i.e 108 jokes into 3 categories), and combinations. In total there were 43 grimace-sound combinations and 2052 grimace-joke combinations	$\mathcal{A} = \{\text{pitch (increase, decrease), rate(increase, decrease), volume(increase, decrease), pause (increase,decrease), } \textit{nop} \}$, where <i>nop</i> is an action that does not change the current state and it is intended to be used when the optimal setup was found. An additional note is that if a marker of the setup (state) is already at maximum or minimum value, it can no longer be increased or decrease.
Reward	<p>set \mathcal{F}_1 and set \mathcal{F}_2 include all occurring relevant smile and laughter events respectively between s_t and s_{t+1}. The reward function was designed as:</p> $R(s_t, a, s_{t+1}) = \frac{1}{2}\mathbb{E}[\mathcal{F}_1] + \frac{1}{2}\mathbb{E}[\mathcal{F}_2]$	<p>Since smiles occur more frequently than laughter, they were more weighted in the reward</p> $R_t = \frac{3}{4} \cdot \mathbb{E}_{\text{smile}} + \frac{1}{4} \cdot \mathbb{E}_{\text{laughter}}$ <p>where $\mathbb{E}_{\text{smile}}$ and $\mathbb{E}_{\text{laughter}}$ are the average probabilities of smiles and laughter detected after the joke of the robot and $R_t : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$ at time step t</p>

On the other hand, Mitsunaga et al. [117] integrated multi-modal social signals directly in the reward function with a Policy Gradient RL to get a Behavior Adaption System. The inputs to the system are focused on the *human's gazing at robot's face* and *human movement distance* as indicators of human's comfort and discomfort. The Behavior Adaption System includes 6 parameters to be adapted:

- The Interaction Distances, based on Hall's definitions [65]:
 1. Intimate Distance
 2. Personal Distance
 3. Social Distance
- The extent to which robot would meet the human's gaze
 4. Gaze-meeting Ratio
- Waiting time between utterance and gesture
 5. Waiting Time
- Speed at which the gestures are performed
 6. Motion Speed

The reward function used was as shown in Equation 3.9:

$$R = -0.2 \times [D_m] + 500 \times \left[\frac{\Delta t_{looking}}{\Delta t_{interaction}} \right] \quad (3.9)$$

where D_m was the movement distance (mm), $\Delta t_{looking}$ was the time human spent looking at robot, and $\Delta t_{interaction}$ time spent for the interaction behavior. The D_m was the sum of the distances that the human moved in all the sampling periods of 200 ms of a behavior and $\left\{ \left[\frac{\Delta t_{looking}}{\Delta t_{interaction}} \right] \times 100 \right\}$ was the % of time the human's face was turned towards the robot with an allowance of $\pm 10^\circ$. The reinforcement method used was PGRL, which directly adjusts the policy without calculating action value functions [163]. The model was evaluated with 15 participants and for most of the users, some of the parameters converged to human preference after 15 to 20 minutes. In addition, it was found that there were some unsuccessful cases and mainly caused by current parameters that were further away from the preferred values and as such the behaviors of the participants did not make a difference for the policies. Consequently, this study shows the possible need of training in simulation prior to interactions with users.

Table 3.8 Summary of Proposed Empathy RL Model by Leite et al. (2012) [105]

Objective	Adaptive Empathy Model to maximize positive feeling after interaction
Method	Multi-armed Bandit problem (for RL policy for selecting empathetic strategies)
States	4 Empathetic Strategies
Actions	<p>Selecting next Strategy</p> <ul style="list-style-type: none"> • If not all strategies were initialised yet: <ul style="list-style-type: none"> * select each strategy by turn • Once all strategies were chosen at least once" <ul style="list-style-type: none"> * the selection changes to selecting a strategy that maximises: $\bar{x}_i + \sqrt{\frac{2 \ln n}{n_i}}$ * \bar{x}_i is the average reward obtained when selecting strategy i * n_i is the number of times strategy i was selected to that point * n is the total number of strategies selected to that point
Reward	$R_i = P(positive)_{after} - P(positive)_{before}$ where $P(positive)$ is the probability of the user feeling positive a few seconds after he/she made a move and before the move was made respectively

Leite et al. [105] presented a multimodal framework to model a user's affective state using visual and task-related input in order to adapt the robot's empathetic responses to the child being interacted with. First, to model the user's affective states, the user's valence of feeling was measured. For a positive feeling to be measured, the user's non-verbal behaviours were either looking at the robot or smiling while task-related features either by user winning, getting better at the game, or catching a piece in the chess game. On the other hand, for a

negative feeling to be estimated user would be either looking at the chessboard or elsewhere while task-related aspects were either the user losing, declining performance in the game, or robot catching a piece in the chess game. Second, the empathetic responses were needed at two points during the game: after the user makes a chess move and when the user was expressing negative feelings. The four empathetic strategies were the following: encouraging comments, giving feedback on move made, suggesting a better move, or intentionally playing a bad move. The evaluation study had three conditions: *neutral*, where no empathy strategy was chosen and the robot would only use facial expressions (e.g. act happy when the use is losing), *random empathy*, where random strategies were selected, and *adaptive empathy*, which is portrayed in Table 3.8. After 40 children interacted with the iCat robot for 10 to 15 minutes, they filled out a questionnaire to measure engagement, help, and self-validation. A significant difference was only found in the help measurement with the robots exhibiting *random empathy* and *adaptive empathy* ranked higher than *neutral*.

Human-robot interactions rely on social signals from humans in order to achieve adaptation. However, these social signals are multi-modal in nature and can be dynamic. First, trying to discretize them would not only lead to large state-spaces, which would require long learning periods and cost computationally, it would also lead to possible losses of information. Thus, defining a reward function that can capture the interaction may help solve that problem as presented in this thesis. Second, in the literature work, most of the robot's possible actions to select from do not capture the robot's embodiment through its multi-modal behavioral capabilities. Consequently, there is still a need to investigate combinations of multi-modal behaviors that are task independent and that can influence the quality of the interaction and the robot's social intelligence and this thesis aims at answering that. Third, training RL models for adaptation and personalization through HRI user studies can be very costly. This thesis additionally intends to contribute to the simulation of HRI environments by designing a simulated human's behaviors interacting with a social robot and portraying the effect of the robot's chosen behaviors on the human.

Part II

Contributions

Chapter 4

Robot Gaze Adaptation using Human Proxemics

This chapter was published as a conference paper as [167]:

K. Tatarian, M. Chamoux, A. K. Pandey and M. Chetouani, "Robot Gaze Behavior and Proxemics to Coordinate Conversational Roles in Group Interactions," 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), 2021

This chapter contains:

4.1	Introduction	41
4.2	Background and Definitions	42
4.3	Adaptive System	42
4.3.1	Conversational Role Coordinator Using Human Proxemics	43
4.3.2	Adaptive Robot Gaze Behavior	47
4.4	Study Design	47
4.4.1	Conditions	48
4.4.2	Hypotheses	49
4.4.3	Measurements	49
4.4.4	Experimental Procedure	50
4.5	Results	50
4.6	Discussion	53
4.7	Conclusion	55

4.1 Introduction

Social intelligence is comprised of many factors, such as how an agent acts, speaks, gestures, and even stands. As a first step to understand how the multi-modal behaviors of the human in an interaction can be utilized to adapt the robot's own modalities and as such act as a socially intelligent agent: this chapter focused on how one modality for human's behavior (*proxemics*) allows the robot to estimate an intention and adapt one of it's modalities (*gaze*) accordingly. More specifically, for example, multi-party interactions are a good use case in which proxemics as modality is crucial and adaptation is necessary.

For instance, in his quest to formulate the success of an interaction, Goffman highlighted the notion of footing, which is the concept used to understand the set-up of a conversation, specially in group interactions [61, 60]. Diving into understanding footing, Goffman introduced the concept of face engagements, which is the set of all instances where two or more participants in a situation are jointly maintaining focus on a single mutual activity [60]. He defined *bystanders* as participants who are officially present in a situation, where more than two participants are also in, but are also officially excluded from the encounter and are not engaged. The situation was referred to as a *partially-focused gathering* [60]. However, when there are only two participants in the situation and there is an encounter, the situation is a *fully-focused gathering*. Group dynamics are an integral part of the make-up of interactions yet can be complex: there are different roles participants can play within a group and their respective spatial distance and orientations give an insight into that.

In human robot interaction, a social robot is not an isolated agent but rather an active participant in social interactions and conversations. As such, it needs to be able to identify the different roles of participants in group dynamics. In addition, it should be able to recognize its own role and adapt accordingly, whether as a speaker, listener, or bystander. Thus it is crucial for the robot to be robust to any changes in the group dynamics being formed around it. Building on findings in proxemics and group interactions, this chapter presents an autonomous adaptive robot gaze behavior based on a conversational role coordinator system, which assigns the roles for the participants in the multi-party human-robot interaction. The system takes into account the position and orientation of the participants to estimate the status of all group members around the robot in addition to the status of the robot itself. Based on the statuses assigned to the participants, the robot adapts its gaze behavior. A pilot study was conducted to investigate how adaptive gaze behavior and group understanding impacts the multiparty interactions with the robot through both behavioral and subjective outcomes.

4.2 Background and Definitions

The aim of this section is to briefly introduce key concepts and definitions based on group HRI literature and based on which the system introduced was inspired. The key definitions are shown in Figure 4.1.

- *Transactional Segment*: In Kendon [86], the transactional segment is defined as the individual's space in which they aim their attention and have the capability of manipulating objects. Additionally, the size of this space may differ based on the activity being conducted and as such by moving their head or upper body orientation they can direct their gaze out of it. For an HRI setting, in [151], the transactional segment is defined as the half circle around the forward direction of the person with a radius of 2 meter. This identifies an object in that segment as the person's implicit attention target.
- *F-formation*: also known as the face formation [112], it is established when two or more individuals arrange themselves so that their transactional segments overlay to form a joint transactional space (known as the *o-space*). Individuals would then have equal and exclusive access to this space so F-formation system is made of cooperative spatial and postural behaviors. Additionally, if someone wishes to join the group then the F-formation needs to be reconfigured [86]. Technological intervention on the structure of F-formations showed that people position themselves to have screen visible [77, 111].
- *Front Zone*: In estimating the participation status of a person based on observations of human interactions, [151] defined the *front zone* as the area across an angle of 120 degree from the front of the person. Moreover, it was concluded that there is a perceived obligation to participate in a conversation when people are in each other's front zones.
- *Gaze Zone*: When two people fall in each other's *gaze zones* thus having their gazes meet, there is now an obligation for the participation in a conversation [151].

4.3 Adaptive System

In this work, an adaptive gaze mechanism system for the robot was developed that allows the robot to react and adjust its gaze based on the current estimate of the conversational role played by the participants in the group. The current estimate is calculated based on the proxemics of the humans around the robot. This aims at making the robot robust to changing group dynamics. Robot gaze cues and shifts have the ability to shape participant roles in a group [127, 131]. This chapter introduces the conversational role coordinator algorithm

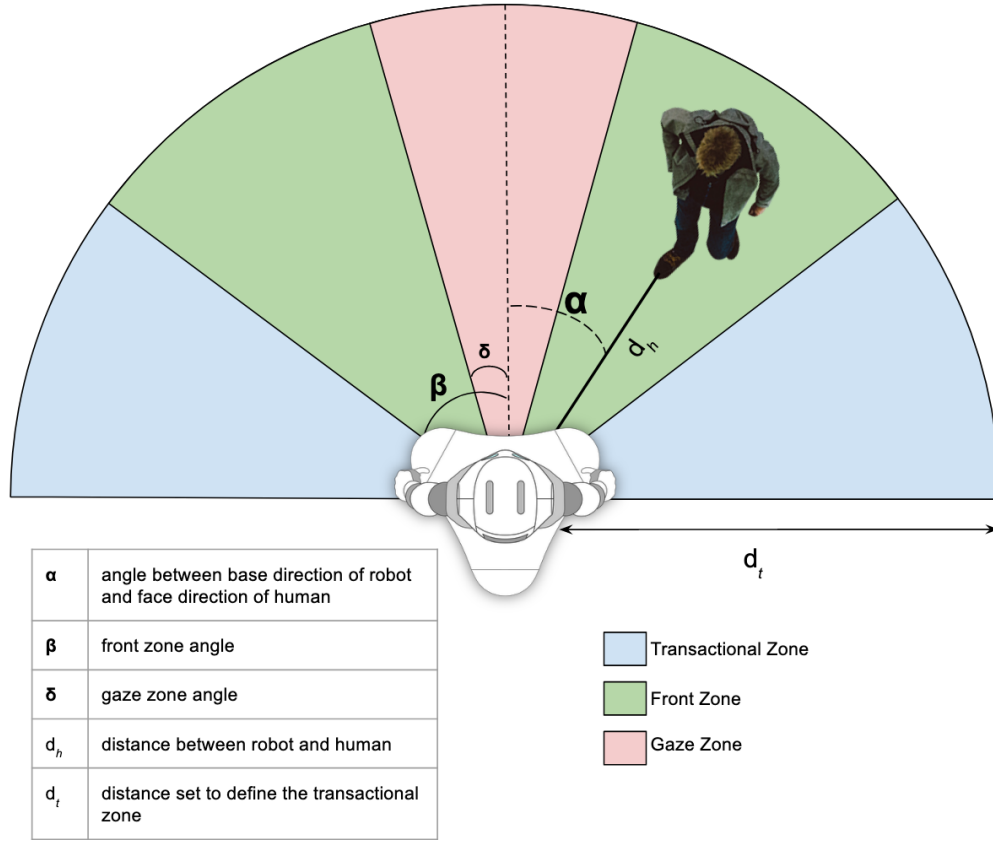


Fig. 4.1 Set-up of Transactional, Front, and Gaze Zones

and the adaptive gaze behavior in the flow of the interaction. It is important to note that this chapter focuses on small group interactions made of three participants in total with the robot and two human participants.

4.3.1 Conversational Role Coordinator Using Human Proxemics

The set-ups parameters, which are shown in Figure 4.1, need to be re-calculated at several instances during the interaction: the distance between the human and robot (d_h) and the angle between the direction of the base of the robot and the direction of the face frame of the human (α). In addition, the remaining parameters need to be set and maintained prior to the interaction and include: β , which is the front zone angle, δ , which is the gaze zone angle, which determine the angular range of the front and gaze zones around the robot respectively, and d_t , which is the distance from the robot that sets the transactional zone.

The robot gathers information needed to evaluate the conversational role status and adapt its gaze behavior accordingly every six seconds, which is a time chosen that give the robot

enough time to naturally finish exhibiting its gaze pattern before re-adapting to changes if detected. The information needed for the robot includes extracting and calculating of the following frames, as shown in Figure 4.2:

- Human Face Frame: the face frame of the humans around the robot
- Robot Base Frame: the base frame the robot
- Robot Gaze Frame: the head frame of the robot in order to adapt it's gaze behavior

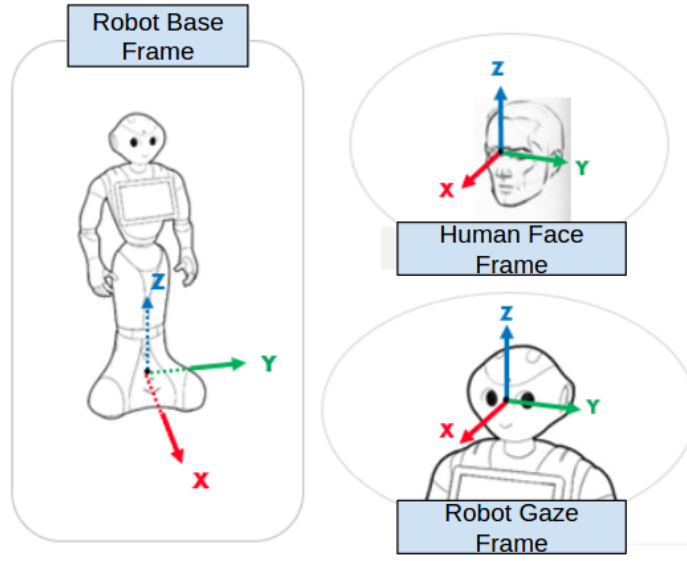


Fig. 4.2 Representation of Human Face Frame, Base Frame of Robot, and Gaze Frame of Robot

The parameters defined in Figure 4.1 were set relative to the base frame of the robot. First, the *transactional zone*, referring to the half circle in front of the robot, had its radius fixed at $d_t = 1.6$. The distance was set less than that of [151] since the social robot used throughout this chapter has a tablet as part of its design and functionality. The tablet must be visible and accessible to the users. Second, the *front zone* defined as the 120° degree fan-shaped area with a 1.6 meters in front of the robot; thus $\beta = 60^\circ$. Third, the *gaze zone* set to a 30° degree fan-shaped area with d_t and $\delta = 15^\circ$. From these frames, d_h and α are calculated for all humans around the robot. *Attention* refers to the head orientation and gaze direction based on the face frame of the human. For example, it describes whether or not the human is looking to their right or left or straight in front of them (thus looking at the robot).

Algorithm 1 shows how the robot evaluates the information input to decide the conversational role status. The three roles that the humans around the robot and the robot itself can

have are: overhearer, bystander, and active. First, if a human was detected to be outside the *transactional zone*, an overhearer role would be assigned since they do not yet show intent of high desire to interact with the robot directly. If all humans detected are overhearers, then the robot is an overhearer as well. Second, if within the *transactional zone*, there is one human alone, then he/she and the robot both have their respective statuses changed to active. Third, if there are multiple people in the *transactional zone*, more examination in proxemics is needed. For all participants within that zone, primary investigation is to check whether equidistant F-formation was formed between participants and the robot, indicating preferred equal access to the conversation, giving the participants in the F-formation an active status. If that is not the case, thus no F-formation was formed, next is the *gaze zone* check, which is within the *transactional zone*. The participant inside the *gaze zone* and the robot gain the active status while all participants outside the gaze zone have the bystander status. However, if there are multiple people in the gaze zone, an attention check is conducted, the participant directly looking at the robot is assigned the active status with the robot while the rest within the *transactional zone* get a bystander status. The algorithm 1 is reran as the perception of the environment by the robot changes. In addition, Figure 4.3 gives an overview on how the robot adapts its gaze mechanism based on the estimation of role coordination algorithm using proxemics of the changing group dynamics.

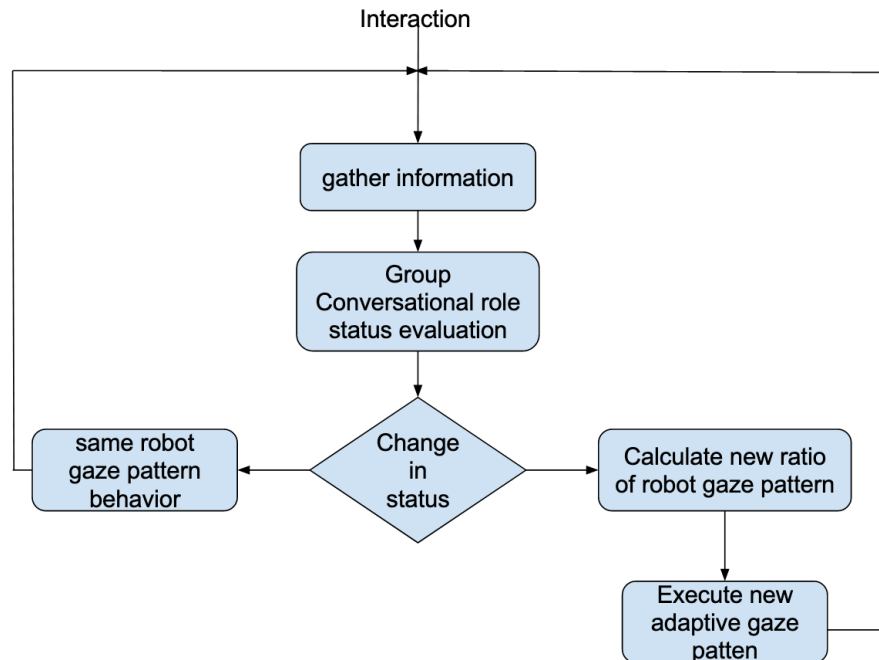


Fig. 4.3 Flow Chart of the adaptive gaze behavior mechanism based on the status evaluation of the conversational role coordinator

Algorithm 1: Group of 3 Conversational Role Decision Making

Data: $[dh_1, dh_2]$ = distances between humans and robot; $[\alpha_1, \alpha_2]$ = angles between robot base and human face frame; d_t = transactional distance; β = front zone angle; δ_1 = gaze zone angle; f-formation = whether or not f-formation was detected; $attention[]$ = attention direction of participants

Result: $status = [status_1, status_2, status_r]$: status of participants and robot

```

while  $i < n$  do
  if  $dh[i] > d_t$  then
     $status[i] = \text{overhearer}$  ;
    count ++;
  end
end
if count == 0 then
  if f-formation == true then
    // Equidistant f-formation Check
     $status_1 = status_2 = status_r = \text{active}$ ;
  else if  $|\alpha_i| < \beta$  and  $|\alpha_j| > \beta$  then
    // Inside the Front Zone
    if  $|\alpha_i| < \delta$  and  $|\alpha_j| > \delta$  then
      // Inside the Gaze Zone
      if  $attention_i == \text{towards robot}$  and  $attention_j == \text{away from robot}$  then
        // Attention Check
         $status_i = status_r = \text{active}$  ;
         $status_j = \text{bystander}$  ;
      else
        // Gaze Zone Check
         $status_i = status_r = \text{active}$  ;
         $status_j = \text{bystander}$  ;
      end
    else
      // Front Zone Check
       $status_i = status_r = \text{active}$  ;
       $status_j = \text{bystander}$  ;
    end
  else
     $status_i = status_j = status_r = \text{active}$ 
  end
else
  if count == 1 then
    // Transactional Zone Check
    while  $i < n$  do
      if  $status[i] \neq \text{overhearer}$  then
         $status_i = \text{active}$ ;
         $status_r = \text{active}$ ;
      end
    end
  else
    // all are overhearers
     $status_r = \text{overhearer}$ ;
  end

```

4.3.2 Adaptive Robot Gaze Behavior

Once the status of the participants and the robot are set, they are used to have the gaze behavior of the robot adapted accordingly, as seen in the flowchart Figure 4.3. Most social robots by design lack articulated eyes, thus head rotations are implemented in coordination with intended gaze movement. In addition, eye gaze alone may not be sufficient to be noticed by participants emphasizing the significance of coordination of head movement and gaze [176]. When the robot has an overhearer status, no direct engagement gaze behavior is activated. The robot instead keeps running algorithm 1 in the background until an active participant is present in the interaction. Once the robot's status is turned to active, an adaptive gaze mechanism is animated. Participants with an overhearer status are ignored by the robot. However, the adaptive gaze behavior is equally split between active participants. For instance, if there are two active participants then a pattern of equal time gazing at each active participant is formed. On the other hand, bystander participants are also recognized by the gaze behavior of the robot but not allocated the same amount of time or attention as the active speaker. The robot glances at the bystander for one third of the amount time it gives to the active speakers. This short amount of time is intended to show the bystander participants that the robot acknowledged his/her presence. For instance, if there is one active and one bystander human in the group interactive around the robot, the robot would form an adaptive gaze behavior pattern, in which six seconds are spent engaged with the active speaker and two seconds to glance at the bystander, the pattern is repeated until a change of status is detected.

4.4 Study Design

In order to evaluate the effects of the conversational role coordinator system using adaptive gaze mechanism and proxemics have in a small group interaction with the robot, a pilot mixed study was conducted with two conditions. The study employed a 2×3 independent groups design. In total, $N = 16$ participants (average age 27) were recruited in 8 pairs. Two participants were needed for each trial per condition thus forming a group of 3 with the robot. The participants were instructed to play a trivia game with the robot made of 3 levels: 1 two player level, where both participants could contribute to answering the question, and 2 one player levels, where only one of the two participants could answer. They were free to choose as they please whose turn it would be for the one player levels without informing the robot. In addition, the participants were instructed that they are free to move in the room however they want.

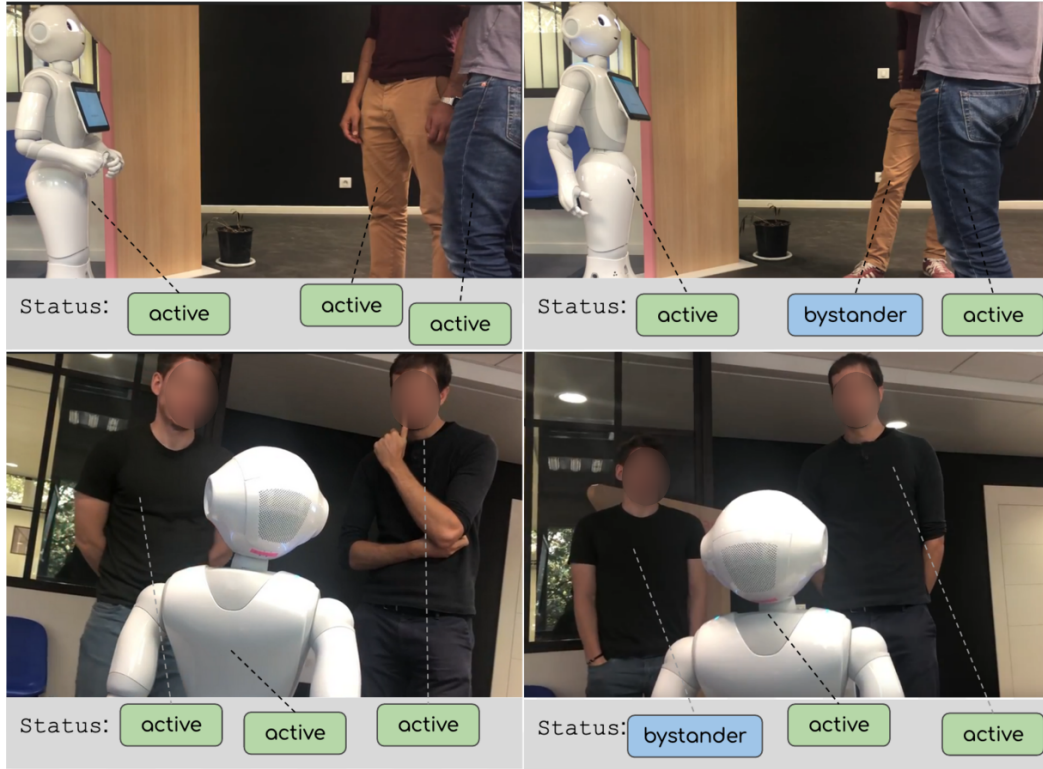


Fig. 4.4 Overview of two sets of interactions between the robot and participants forming a small group. The statuses represent each conversational role for each participant estimated by robot in different contexts and times during the interaction

4.4.1 Conditions

In the first condition, the robot displayed an autonomous and adaptive gaze mechanism using the output statuses of algorithm 1 as discussed in the previous Section 4.3. In the rest of the chapter, this condition and system is referred to as Group Conversational Role Coordinator and shown in Figure 4.4. On the other hand, in the second condition, which is the control condition, the robot was in a *basic awareness behavior*. The *Basic Awareness Behavior* refers to an autonomous ability on the Pepper Robot.¹ Basic Awareness Behavior enables the robot to react to certain stimuli, which include touches on bumpers, touches on hands, or human detection, by looking in the direction of that stimuli using the head rotation of the robot. Thus, the robot would look only at the first person it found until another stimuli would be detected. When no person was detected, the robot would return to a neutral position.

¹Pepper Robot by SoftBank Robotics (<https://www.softbankrobotics.com/emea/en/pepper>)

Did you feel attended to by the robot?
Was the robot giving you enough attention?
Did you feel that the robot was looking at you?
Do you think the robot considered you as an important player?

Table 4.1 Extra Questions: Feeling Attended To (cronbach's alpha = 0.71)

4.4.2 Hypotheses

According to literature review discussed in Chapter 2, it was suggested that a robot's gaze behavior can shape conversational roles [127] and adaptive gaze behavior to group dynamics increase feelings of inclusion, sense of belonging to the robot's group [176], and encourage even participation within the group [59]. Based on the state-of-the-art findings, the following hypotheses were formulated:

- **H1** Participants will stand closer to the robot even when playing a bystander role while they interact with the robot having the adaptive gaze mechanism based on the conversational role coordinator.
- **H2** Robot with an adaptive conversational role coordinator will be perceived as more adaptable, enjoyable, sociable, and socially present.
- **H3** Participants will feel more attended to by the robot with the conversational role coordinator than by the robot performing basic awareness behavior.

4.4.3 Measurements

In order to evaluate **H1**, the distances d_h of each human around the robot and the angle α they were forming with the robot, as shown in Figure 4.1, were extracted every six seconds throughout the entire interaction.

Subjective measurements were used to assess **H2** and **H3**. The participants were asked to answer a questionnaire ranked from 1 (strongly disagree) to 5 (strongly agree) and composed of an ALMERE questionnaire and additional questions to evaluate whether participants felt attended to by the robot. The ALMERE questionnaire was designed to assess the acceptance of assistive social agent technology [73]. The questionnaire includes the following constructs: anxiety (ANX), attitude towards technology (ATT), perceived adaptability (PAD), perceived enjoyment (PENJ), perceived sociability (PS), and social presence (SP). The additional questions asked in the questionnaire after the ALMERE questions are presented in Table 4.1.

4.4.4 Experimental Procedure

Each interaction required a small group of two participants at a time to play a trivia game with a Pepper robot. The trivia game consisted of general knowledge questions to which the players had to answer by either 'True' or 'False'. The game was made up of three levels not based on difficulty but rather on whether it was for one player or two players. The robot announced the type of level before asking a set of questions. The first level was for two players and the second and third levels were for one player where the participants had to switch turns amongst each other. It was designed as such to create different scenarios with various possible combinations of statuses such as two active and one active with one bystander. At the beginning, the robot would greet the players and at the end of the game say goodbye and thank them for playing. During all levels, the participants were informed that they were free to move as they wish. For levels two and three, the participants were told that they were free to choose amongst each other which one would be the main player. Thus the robot was ambiguous to the main player decisions made. The robot in the condition of *conversational role coordinator* relied only on the proxemics and Algorithm 1 to predict statuses of participants around the robot and accordingly adapt its gaze behavior, while in the *basic awareness behavior* condition the robot chose the participant to look at based on which of the two was detected first. The participants were randomly assigned to the two conditions. The interactions with the robot took on average 15 minutes. At the end of the entire interaction, the participants were given the questionnaires to answer.

4.5 Results

First, a factorial ANOVA analysis with post-hoc Tukey HSD was done on the distances measured by all participants throughout all three levels of the interaction per condition, shown in Figure 4.5. A significant difference was found between the two conditions, robot with an adaptive conversational role coordinator and robot with a basic awareness behavior, with $F(1, 44) = 7.635$ and $p < 0.01$. The average distance for participants interacting with the basic awareness behavior was $d_h = 1.637, SD = 0.833$; meters while that of the ones interacting with the conversational role coordinator was $d_h = 1.31, SD = 0.328$ meters. In addition, there was a significant difference between the distances of the various roles played by participants in the interaction (a bystander player, a main player, and one of the two main players) with $F(2, 44) = 25.665$ and $p < 0.001$. Furthermore, when playing a bystander role, participants interacting with the Group Conversational Role Coordinator stood closer at ($M = 1.68, SD = 0.11$) meters where as participants stood further away when interacting with the Basic Awareness Behavior at ($M = 2.03, SD = 0.83$) meters. All distances are summarized in Table 4.2.

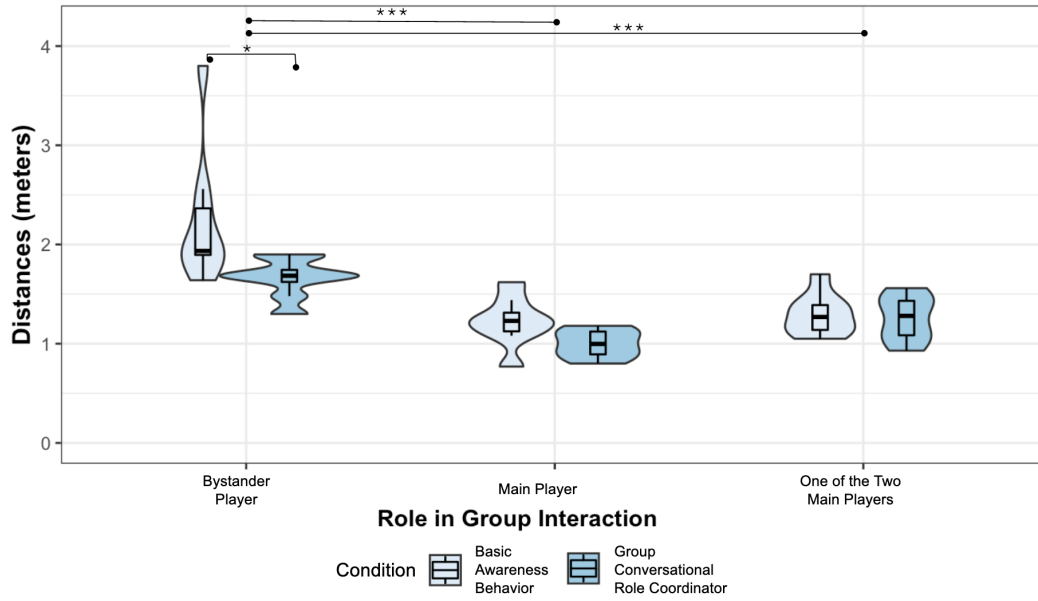


Fig. 4.5 Distances (meters) maintained by participants per condition and per role they played during the interaction: when they were playing a bystander role, main player role, and one of the two players. There was a (**) significance, for $p < 0.01$, between the two conditions. In addition, (*) represents significance of $p < 0.05$ between the two conditions for the Bystander Player and (***) represents significance of $p < 0.0001$ between different roles in the group interaction.

<i>Distances (in meters) of Participants around the Robot</i>						
Roles	Group Coordinator		Basic Awareness		Total	
	M	SD	M	SD	M	SD
One of Two Main Players	1.25	0.23	1.29	0.2	1.27	0.21
Main/Active Player	0.99	0.13	1.22	0.25	1.11	0.23
Bystander Player	1.68	0.11	2.39	1.09	2.03	0.83
	1.31	0.32	1.63	0.83	1.47	0.64

Table 4.2 Summary of average distances held by participants throughout the different roles in the interaction per condition and in total

Second, taking a closer look at the angle α , defined in Figure 4.1, done by group members around the robot, a factorial ANOVA with a post-hoc Tukey was performed, shown in Figure 4.6. No significant differences were found between the two conditions, robot with Basic Awareness behavior and Group Conversational Role Coordination, nor between the different roles participants played throughout the interaction with $F(1, 44) = 0.991, p > 0.05$ and $F(2, 44) = 0.107, p > 0.05$ respectively. Additionally, when the interaction was at two

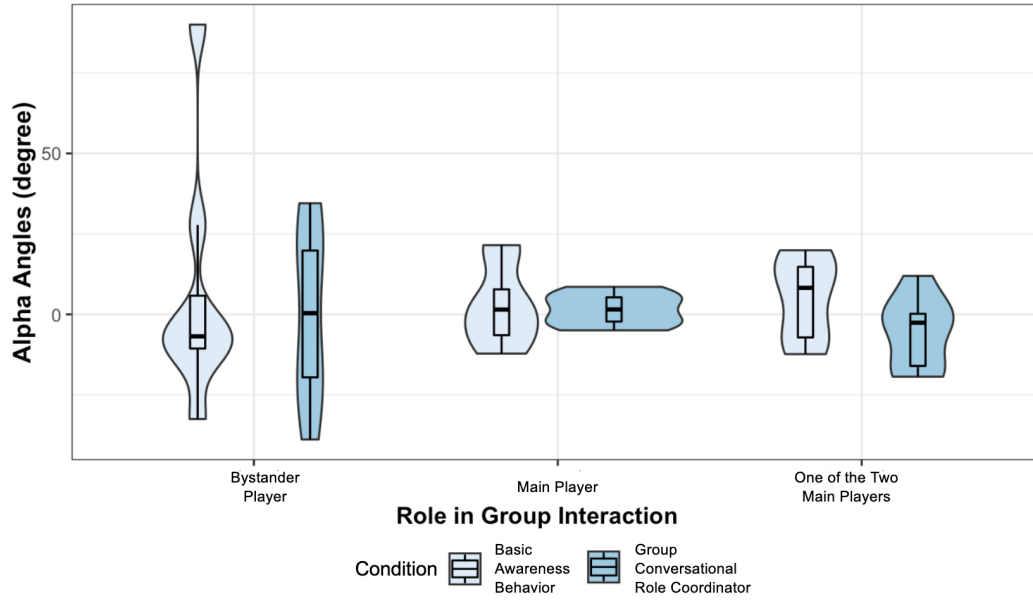


Fig. 4.6 α Angles (degree) maintained by participants per condition and per role they played during the interaction: when they were playing a bystander role, main player role, and one of the two players. Alpha angle refers to the angle define in Figure 4.1

players level, the angle α was smaller when interacting with the robot with the conversational role coordinator than with that with basic awareness; thus participants were more aligned with the facing direction of the robot. With the findings based on distances and angle α of participants, **H1** was partially supported.

Third, regarding the subjective measurements, a Two-way ANOVA analysis was conducted for the ALMERE Questionnaire constructs rated by participants, shown in Figure 4.7. A significant difference was found between the two conditions with $F(1, 41) = 4.526, p < 0.05$ with an effect size $\eta^2 = 0.46$. The robot with the Conversational Role coordinator rated significantly higher on the Perceived Adaptability, Perceived Sociability, and Social Presence constructs. This supported **H2**.

Fourth, the additional questions asked to better understand whether or not participants felt attended to by the robot had a Cronbach's $\alpha = 0.71$. Accordingly, the combined questions presented as one construct in Figure 4.8. An One-way ANOVA with pairwise comparison analysis revealed a significant difference between the two conditions, robot with a group conversational role coordinator versus a robot with a basic awareness behavior, with $F(1, 41) = 4.526, p < 0.05$. The participants rated the questions higher after having interacted with the group conversational role coordinator robot than with the robot with basic awareness behavior. **H3** was supported.

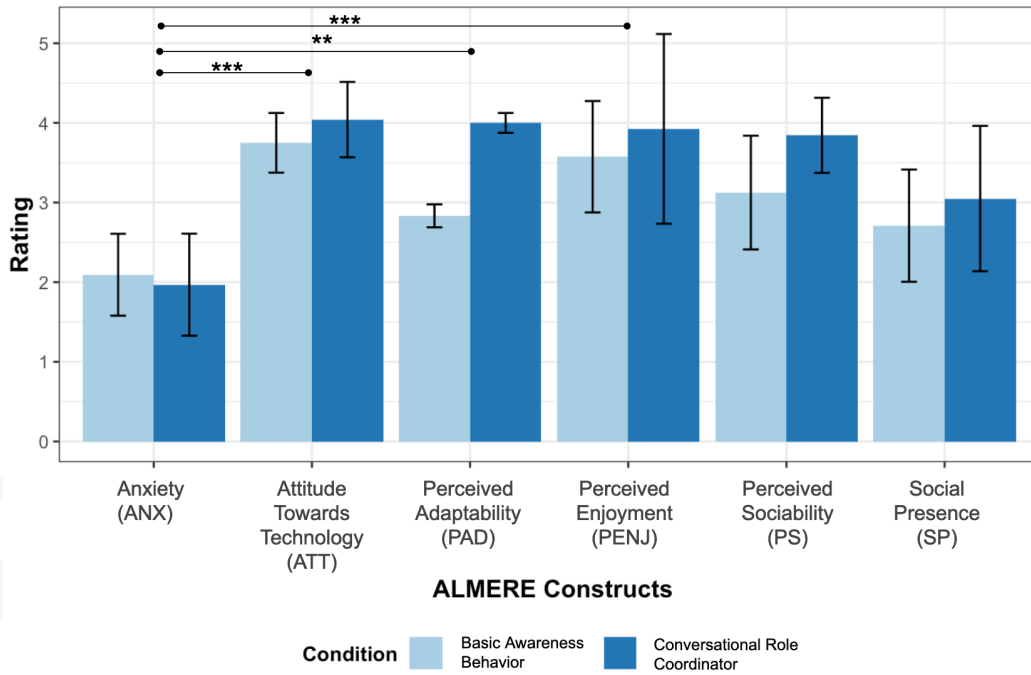


Fig. 4.7 Ratings of ALMERE Constructs per condition by participants with (*) significance for $p < 0.05$ between the two condition. (**) refers to $p < 0.01$ and (***) to $p < 0.001$

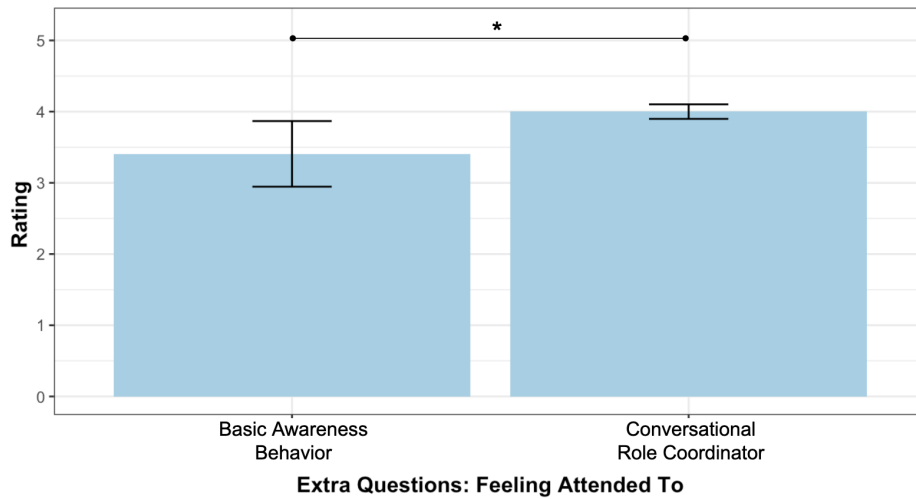


Fig. 4.8 Ratings from extra questions related to 'Feeling Attended To' per condition by participants with (*) ($p < 0.05$) significance between the two conditions

4.6 Discussion

While this was a primary study to get an insight on the subjective and behavioral outcomes of having an adaptive gaze behavior for group conversational role coordinator, there were

leading primary findings. First, the distances of participants in the small groups being formed with the robot showed that people stood significantly closer to the robot when it performed group conversational role coordination through adapted gaze behavior based on the roles they were playing with the group, partially supporting **H1**. When both members of the group were active participants/ main players, they stood at similar distances away from the robot. However, when the level switched to one player, the main/active player moved even closer to the robot when interacting with an adaptive group gaze behavior robot. On the other hand, the bystander player moved much further away when interacting with the basic awareness robot than with the adaptive group coordination robot with a difference between the two averages of $\Delta d_h = 0.7m$. This relative big difference may be due to the bystander participant thinking they needed to move further away to not disrupt the interaction between the main player and the robot specially if the robot was prior gazing at the bystander player. On the other hand, in the adaptive group behaving robot, the bystander participants felt more comfortable staying close to the interaction between the main player and the robot since it did not take away attention from the main player. The participant interacting with the conversational role coordinator robot was also able to listen in on the interaction and be acknowledged by the robot.

Even though, there was no significant difference between the two conditions in the α angle, it was smaller in the case when there were two main players in the interaction with the group conversational role coordinator robot. This shows that the members of the small group with the adaptive gaze behavior robot were standing closer to each other and closer to the direction of the robot than in the group with the basic awareness robot. This may be an interesting insight as to how participants form groups with the robot when it is able to give equal adaptive attention to the small group members; however, a bigger study is needed to investigate that.

Second, the subjective measurements presented a significant difference between the two conditions. Through the ALMERE questionnaires ratings, **H2** was supported. Participants interacting with the robot having an adaptive conversational role coordinator was rated higher on perceived adaptability, perceived enjoyments, perceived sociability, and social presence. Third, the ratings on the extra questions asked addressing the perception of being attended to by the robot were significantly different between the two condition supporting **H3**. Furthermore, in Figure 4.8, there was a larger standard deviation and variation in the basic awareness behavior condition. This may be due to the fact that within that condition and because of the behavior of the basic awareness, one of the two members of the small group got the most engagement and attention from the robot while the other was excluded from the gaze direction of the robot even when they played the main player role. On the other hand, the conversation role coordinator gave relatively equal attention to both players.

Furthermore, it is also important to note the limitations of this study and its findings. Group dynamics are very complex in nature and involve multiple nonverbal and verbal modalities to be taken into account including culture and personalities of the participants [13, 82]. The system presented in the chapter is a primary and simplified approach to try and allow the robot to estimate the conversational roles. Algorithm 1 depended on inputs from the sensors of the Pepper robot used. Technical limitations in the application may have included false positives of human positions detected by the robot and/or noise in the signals. Moreover, this study is a primary pilot study and perhaps a greater number of groups with more participants or more robot need to be explored.

4.7 Conclusion

Prior work has shown that proxemics is a crucial tool in understanding group formations and dynamics and gaze can be an adaptive tool in group interactions. Building on prior work, this chapter introduces an autonomous conversational role coordinator system that assigns to the group participants around the robot three different statuses, which are active, bystander, and overhearer, based on their proxemics, including distance and orientation. Once the statuses are assigned, the robot readjusts its gaze pattern in order to adapt to the group dynamics giving active participants equal gaze attention, bystanders simple acknowledgements, and ignoring overhearers. This system was evaluated through a pilot study where participants played a trivia game with the robot and had different roles within the interaction. Primary results have shown that participants interacting with a robot having this system are more likely to stand closer to the robot, perceive the robot as more adaptable, sociable, and socially present as well as feel more attended to by the robot. These findings may lead to a better understanding of group interactions with social robots as they become more present in everyday life, such as hospitals, retail stores, schools, and perhaps even homes, thus ushering the way for more socially intelligent interactions.

This Chapter gave an initial insight into how the behaviors of humans can be utilized by the robot to adapt its own behavior and appear as a socially intelligent agent. This was done by only looking at one modality exhibited by the human and one modality adapted by the robot. The primarily findings from the pilot study hint into promising results for further development. The next Chapter builds on this and investigates how each modality making up the multi-modal behavior of the robot can be designed and implemented and studies the extend of their effects and impacts on an interaction, specially the behavioral outcomes of the human.

Chapter 5

Investigating the Synthesis & Effects of Multi-Modal Robot Behavior

This chapter was published as a journal paper as [168]:

Tatarian, K., Stower, R., Rudaz, D. et al. How does Modality Matter? Investigating the Synthesis and Effects of Multi-modal Robot Behavior on Social Intelligence. Int J of Soc Robotics (2021).

This chapter contains:

5.1	Introduction	57
5.2	Implemented System for Multi-Modal Behavior	58
5.2.1	Social Gaze Mechanisms	59
5.2.2	Gestures	62
5.2.3	Proxemics for Approaching Human	62
5.3	Design Method and Evaluation	63
5.3.1	Social Dialogue	64
5.3.2	Design & Materials	65
5.3.3	Participants & Procedure	65
5.3.4	Hypotheses	67
5.4	Results	68
5.4.1	Distances from Pepper	68
5.4.2	Self-Disclosure	70

5.4.3	Accepting Water Offered	71
5.4.4	Social Behavior	72
5.4.5	Liking	77
5.4.6	Voice Recognition Errors	80
5.5	Discussion	80
5.6	Conclusion	83

5.1 Introduction

By merely observing humans, one can directly infer that no social interaction takes place without cues, whether verbal or nonverbal, that allow others to interpret behaviors and reasonably estimate intentions [147]. Furthermore, those verbal and nonverbal cues have an effect on others by eliciting tangible change in their observable behavior or even internal changes, e.g., awareness of a particular social setting [179]. Moreover, proper communication and exchange of information is crucial to a human's need to feel connected, promote well-being, and gain acceptance by social groups [144]. However, these powerful social signals and nonverbal behaviors are complex and multi-modal. They are made of different combinations of modalities and cues such as *kinesics* (e.g., gestures) [97, 48], *gaze behavior* [85], and *proxemics* (e.g., management of space and environment) [68]. Similarly, these multi-modal nonverbal behaviors hold several functions, which include the ability to understand and manage others in social interactions and "act wisely in human relations", and as such contribute to one's *social intelligence*. [169].

In today's world, humans not only have to interact with each other, but also with machines, including robots. With robots gaining further presence in a human's everyday life, synthesizing and understanding these multi-modal behaviors is crucial to designing better and more appropriate human-robot interactions. In an attempt to solve this issue, some studies have been inspired by human-human interaction to design rule-based algorithms targeted at investigating individual modalities [11, 151, 109]. In contrast, other research has focused on data-driven learning methods designed to synthesize multi-modal behavior, however, lacking a clear understanding of the effect of each modality forming the multi-modal behavior [101, 135]. Thus, there is still a need to investigate how such modalities can be combined and the effect and function of each when performed in a multi-modal combination. This will allow for a better understanding of how and when the robot could use combinations of different modalities to appear as a socially intelligent agent and express intentions and information using verbal and nonverbal behavior more naturally.

This chapter presents a system of multi-modal behaviors comprised of the following modalities: gaze, kinesics, proxemics, and social dialogue. The system was evaluated objectively by studying the behavioral outcomes. In addition, critical to evaluating peoples interactions with social robots is also the extent to which they like the robot, or form a general positive impression during their interaction. Thus, in this study a subjective measure of liking was also included through a self-report "liking" scale. Instead of analyzing the effects of each modality by contrasting them in isolation and thus losing possible coupling effects, this chapter compares a version with all implemented modalities together with versions, in which each modality is subtracted in turn.

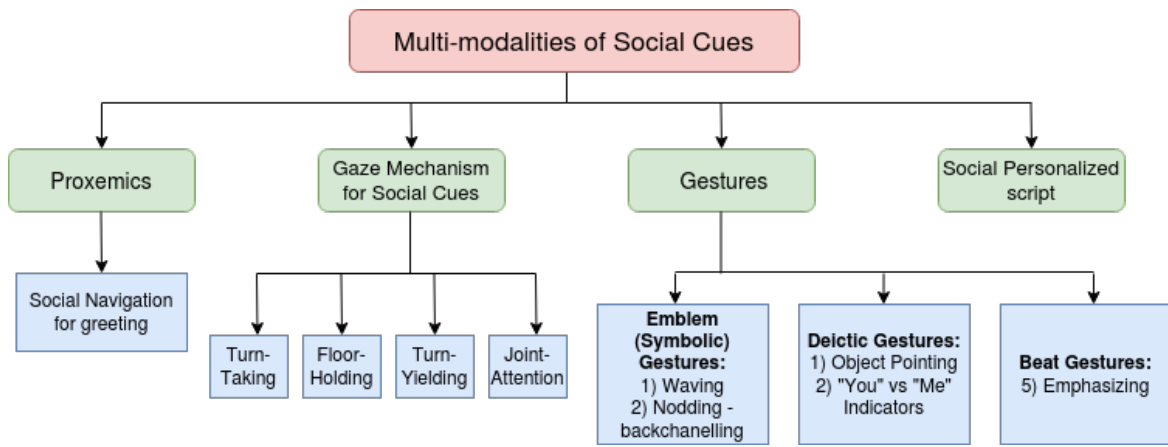


Fig. 5.1 Summary of Multi-Modal Social Cues

5.2 Implemented System for Multi-Modal Behavior

This section introduces the implemented system to achieve multi-modal behavior on the Pepper robot. However, the system can also be implemented on other social robots. The source code for the entire system has been made available online¹. The overall scheme of the multi-modal social cues synthesized are shown in Figure 5.1. The system is composed of the following modalities: proxemics, gaze mechanism, gestures, and a social dialogue. A sample extract of the system implemented is shown in Figure 5.2.

¹Multi-modal Social Cues System Implementation GitHub Repository https://github.com/KarenTatarian/multimodal_socialcues

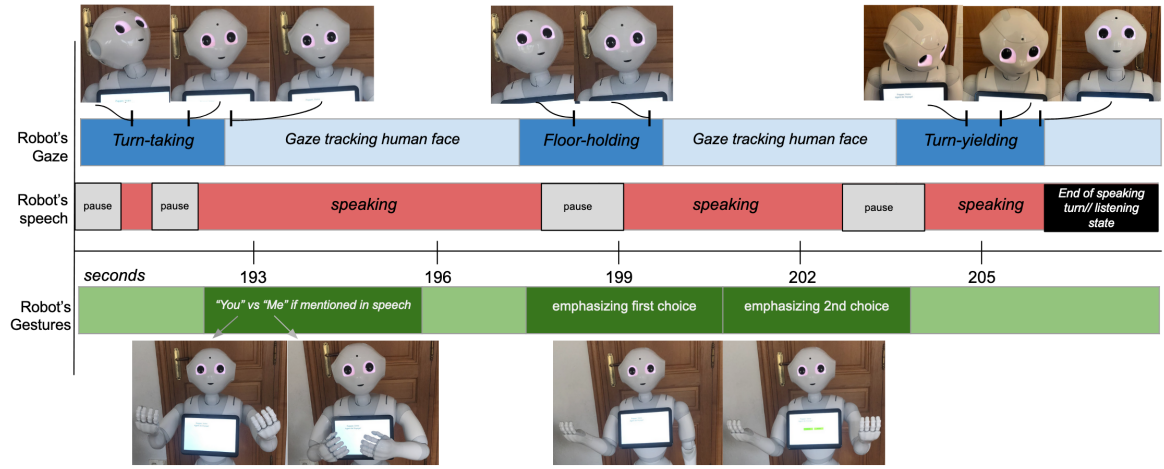


Fig. 5.2 Sample of time-line including speech, gaze mechanisms (turn-taking, floor-holding, turn-yielding), and social gestures (deictic gestures: "You" vs "Me" if mentioned in speech, beat gesture: emphasizing the two choices user needs to select from)

5.2.1 Social Gaze Mechanisms

Since several humanoid robots lack expressive eyes that can be controlled, the presented social gaze aversions are achieved using head motion control. The social gaze mechanisms presented here were designed and implemented to fulfill the following functions: joint attention, turn-taking, floor-holding, and turn-yielding. When not performing these gaze aversions, the robot would be gaze tracking the human it is interacting with at all times. For this reason, information about the human can be extracted, notably the 3-dimensional frames of the human's face and the robot's gaze. This allows the robot to carry out all implemented gaze-averted head motions on the robot with respect to the frame of the human's face. Thus the design of each social gaze head movement was a combination of dynamics, magnitude, and duration, all of which are crucial for the social gaze motion to achieve its function naturally. A summary of the gaze mechanisms can be found in Figure 5.3.

Gaze aversion for turn-taking in human-human interaction as well as human-robot interaction holds a cognitive function; it gives the speaker more time to better plan and address their speech while also avoiding possible external extractions [15, 11]. For this reason, the turn-taking gaze behavior was given the relatively longer duration of 2.5 seconds. As for the floor management and turn-yielding functions, which take place during and at the of the speaking turn, they were assigned a shorter duration of 1.5 and 1.2 seconds each respectively. The longest duration was designed for the joint attention gaze, which has the duration of 3.8 seconds and the function of indicating and referring to an object of discussion. The angle rotations, duration length, and directions of the gaze aversions were selected and designed

based on the gaze aversion system, with similar functionalities, implemented on the NAO robot (SoftBank Robotics) in [11], which are as well based on the findings of Kendon [86] for gaze aversions in human-human interaction. In turn, these variables were tuned to be suitable for the robot Pepper. Examples of how the duration of each gaze mechanism was synthesized is shown in Figure 5.2.

Functions	Duration (seconds)	Direction	Magnitude
Joint Attention	3.8	Right	Head Yaw of -46° & Pitch of -4.6°
Turn-taking	2.5	Upward, left-sided	Rotation $\theta = 33.7^\circ$ in the (y-z) plane of gaze frame
Floor holding	1.2		
Turn- Yielding	1.5	Downward, Left-sided	$x'' = [(1-\cos\theta)]*x + y*\sin\theta - \delta_1$ $y'' = -\sin\theta*x + (1-\cos\theta)*y$ $z'' = \delta_2$ $\theta = 12.318^\circ$, $\delta_1 = 1.2$, and $\delta_2 = -0.6$ $x, y, \& z$ are the coordinates of human head frame

Fig. 5.3 Summary of gaze mechanisms

The goal is to generate head motions, which serve as social gaze cues, and to realize a natural and subtle appearance while still fulfilling their purpose. The directions of the head rotations performed were chosen based on human-human as well as HRI studies done to better understand gaze aversions during interactions [11, 85, 125]. For the turn-taking and floor-holding cues, which signify that the robot is in a cognitive phase and holding the speaking role, a vertical head movement was implemented. The gaze frame of the robot is the frame it produces as it is tracking the human's face as seen in Figure 5.3, thus the gaze frame is aligned with the human's face frame. At each instance in which the turn-taking and floor-holding gaze mechanisms were called, the gaze frame performs a rotation of 33.7 degrees in the (y-z) plane with respect to the frame on the human's face resulting in an upward slightly left-sided rotation. As for the turn-yielding gaze, in order to achieve a smoother

and more subtle gaze to indicate the shift in speaking roles and after manually fine tuning, Equation 5.1 represents the translations implemented for the (x, y, z) of the robot's gaze frame respectively

$$\begin{aligned} x'' &= [(1 - \cos \theta)x + \sin \theta y - \delta_1] \\ y'' &= -\sin \theta x + (1 - \cos \theta)y \\ z'' &= \delta_2 \end{aligned} \quad (5.1)$$

where $\theta = 12.318^\circ$, $\delta_1 = 1.2$, and $\delta_2 = -0.6$ and $x, y, \&z$ are the coordinates of the frame of the human's face. This resulted in a downward left-sided rotation as seen in Figure 5.4, which illustrates an example of the execution. Finally, the head rotation for joint attention was not intended to be subtle but rather draw the attention of the user towards the object it is referring to. The head rotation was done with a head yaw of -46° and a head pitch of -4.65° with respect to the object it is referring to and with a Bézier curve velocity profile for smoothness in the animation. This resulted in an animation that starts slow, speeds up during the main shift, and finally slows down before it ends. For instance, for the gaze mechanisms implemented as head-motions, the head of the robot would start rotating slowly then rapidly towards target frame then slowly returns back to its initial gaze frame as shown in the example of Figure 5.4.

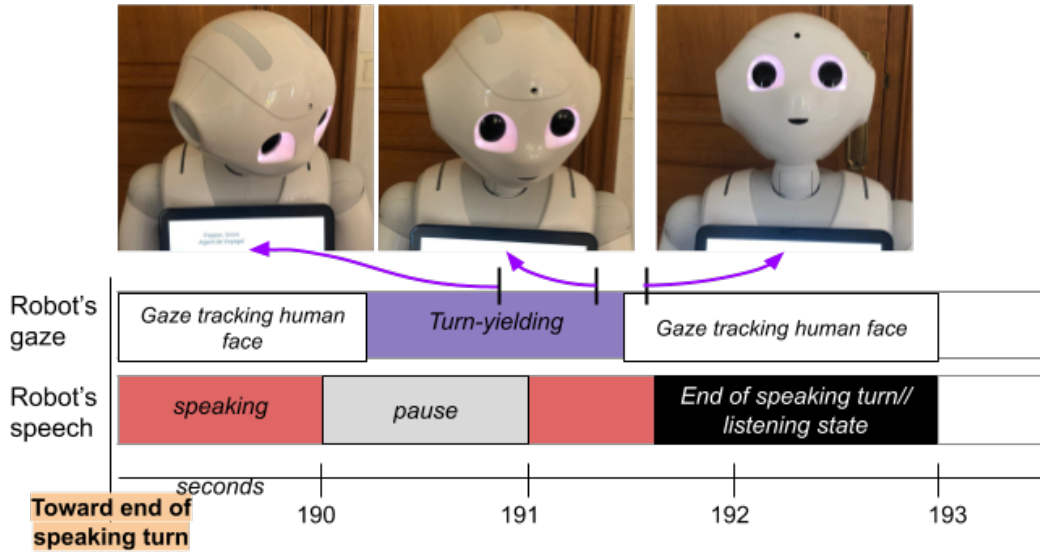


Fig. 5.4 Turn Yield Gaze Mechanism

5.2.2 Gestures

Three of the five categories of gestures classified in [113, 97], are implemented in this multi-modal system through five gestural designs. First, in the emblem category, two gestures were implemented: the *wave gestures* for greeting users and closing interaction and *nodding* as back-channeling. When the robot enters a listening state, it nods in response to additional information the user provides while speaking. Whenever a new sentence is detected by the robot, it executes a back-channeling nod animation and waits for a next sentence to be heard. The back-channeling nod animation consists of three consecutive nods over a span of 2.5 seconds. Second, in the deictic category, two types were represented: *object pointing gesture*, which accompanied the joint attention gaze and performed as a right arm extension, and *self/other references*, which are the “you” vs “me” indicators, as portrayed in Figure 5.2. In the latter, the “you” gesture was designed as both of the arms open and extended from the robot’s body and pointing upwards towards the human with the open palms. The “you” gesture accompanied questions, in which the robot infers about the user. On the other hand, the self-pointing or the “me” gesture has the robot’s arms closed and directed towards the body and it is used by the robot to introduce itself, its tablet modality, and state an opinion. Furthermore, beat gestures were used in the question series to emphasize the options. In order to generate smooth behavioral social cues, Bézier curves were used to implement the velocity changes of all gestures.

5.2.3 Proxemics for Approaching Human

According to the findings in [156, 99], the most optimal and socially acceptable path to navigate towards a static standing human is in a straight line in the human’s sight zone while maintaining a minimum distance of about 0.5 meters (or 1.5 meters for sitting humans). This is due to safety and visibility constraints to minimize human discomfort. In addition, approaching the human for initiating an interaction was ranked as most appropriate when the robot navigated in the human’s front zone (120° cone-shaped area in front of a person’s head) or his/her gaze zone (30°) [151]. For this system, once the human has been detected at the beginning of the interaction, the robot would first greet the human by speech and a wave gesture before beginning navigation towards the human. The default speed of the robot Pepper is 0.35 m/s. However to avoid a recoil movement by the human seen in some user experience testing, the speed for greeting navigation was slowed down to 0.25 m/s. The robot would then navigate to establish a distance of 0.85 meters between itself and the user. The robot would navigate while maintaining gaze directed towards the participant. The distance of 0.85 meters was chosen for four main reasons. First, this distance allows the robot to continuously track the human’s face regardless of his/her height. Second, the human at

this distance is able to clearly see the different gaze and gesture behaviors generated by the robot. Third, this distance eases the access and view of the tablet on the robot. Fourth, 0.85 meters is still within Hall's defined personal distance, in which friendly interactions take place [65]. Once the desired distance has been reached, whether by navigation of the robot or chosen distance by human, the start button pops up on the tablet to continue the rest of the interaction.

5.3 Design Method and Evaluation

In order to directly compare the effects the different modalities of the robot's behavior have on the user's behavior and attitude, the chosen scenario for the interaction was planning for a hypothetical holiday. The robot acted as a travel agent helping the user plan his/her next vacation and it exhibited one of five behavioral conditions:

- Multi-modal Interaction, which is all modalities including social dialogue, (*Social Gaze + Gestures + Proxemics + Social Dialogue*)
- Minus Proxemics, which is all the modalities excluding proxemics, (*Social Gaze + Gestures + Social Dialogue*)
- Minus Social Dialogue, which is all the modalities excluding social dialogue (*Social Gaze + Gestures + Proxemics*)
- Minus Gestures, which is all the modalities excluding social gestures, (*Social Gaze + Proxemics + Social Dialogue*)
- Minus social gaze, which is all the modalities excluding social gaze, (*Gestures + Proxemics + Social Dialogue*)

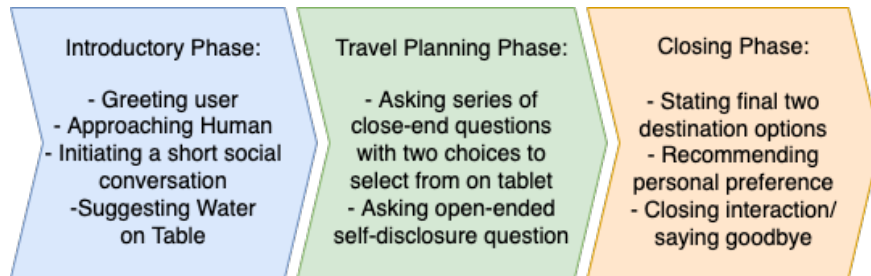


Fig. 5.5 Summary of flow of interaction

The flow of interaction went as seen in Figure 5.5: first in the introductory phase, the robot greeted the user (highlighting the gesture modality), approached the human until the

desired distance is established, engaged the user in a short social dialogue and then offered the participant some drinking water, using joint attention gaze and pointing gesture. Second was the travel planning phase, where the robot started asking a series of questions about the travel and vacation preferences, followed by a self-disclosure segment, where an open ended question was asked to know more about the user and the robot entered a listening state and demonstrated back-channeling. The third and final part was the closing phase, where the robot suggested two options based on answers previously provided by participants and recommended its personal preference between the two. Once the user made a final decision on a travel destination, the robot concluded the interaction and waved goodbye.

	Conditions	
	Multi-modal Interaction, Minus Proxemics, Minus Gestures, and Minus Social Gaze	Minus Social Dialogue
Openings/Closing	Social small talk <i>ex: "How are you today?" ... "great"</i> <i>ex: "I am so happy to meet you!"</i>	Formal talk <i>ex: "Is it the first time you come here?"</i> <i>ex: "You are the fourth person today whom I will help plan their vacation"</i>
Replies	Personal preference replies <i>ex: "Excellent choice, I also like this location!"</i> <i>ex: "I also find it awesome to travel by train because it's much more comfortable!"</i>	Non-personal general replies <i>ex: "Many people like this location"</i> <i>ex: "Traveling by train is more comfortable"</i>

Table 5.1 Social dialogue designs per condition

5.3.1 Social Dialogue

The dialogue throughout all conditions was adaptive: the robot's answers depended on what the user's previous choices were. In addition, since it was a travel planning scenario, all replies from the robot were consistent with the decisions the user made. For instance, if the person selected that they prefer to travel by train then the robot would suggest a destination that can be reached from Paris by train (e.g., Amsterdam) and similarly for the other descriptions, e.g., city or beach, solo trip or with friends and loved ones, culture or activities etc. However, the social dialogue modality differs in the social and friendly openings and replies as summarized in Table 5.1.

5.3.2 Design & Materials

An independent groups design was used, with the independent variable being multi-modal behavior with 5 levels: first: multi-modal interaction referring to all implemented modalities (proxemics, social gaze, gestures, and social dialogue), second: minus proxemics - referring to all implemented modalities except for proxemics, third: minus social dialogue - referring to all implemented modalities except for social dialogue, fourth: minus gestures - referring to all implemented modalities except for gestures, and fifth: minus social gaze - referring to all implemented modalities except for social gaze mechanisms.

The dependent variables were extracted using recorded logs from the robot application, data extracted from the recorded videos, and self-report questionnaires. First, the logs provided from the robot application include information on the position of the human relative to the robot extracted every 5 seconds as well as the angular facial frame information, which were extracted before the execution of every social gaze aversion. In addition, the time it took the user to press the buttons on the tablet and to take decisions as well as the decisions made were recorded.

Second, the videos recorded were used to annotate and obtain the verbal and nonverbal responses and behaviors of the user throughout the interaction, including if the user accepted the water offer, back-channels performed, verbal responses, total speaking time, amount of information shared, number of audio/voice recognition errors that may have occurred, and gestures performed. It is critical to note that the back-channels of the users in this set-up refer to both non-lexical back-channels, such as “uhh”, “yeah”, “mmm”, .. etc., phrasal back-channels, such as “wow”, “great”, .. etc., and gestural back-channels, e.g., nodding. However, facial expressions were not considered.

Third, a self-report questionnaire was used to evaluate the subjective measurement for robot’s Likeability or *liking* and was comprised of 9 items. The participants were asked to rank how well they agree with the statements in Table 5.2 on a scale of 1 to 7, with 1 agree with the least and 7 agree with the most. These items were then averaged to form a scale with good reliability $\alpha = 0.88$. The scale was normally distributed, $W = 0.99, p = 0.35$.

5.3.3 Participants & Procedure

115 participants were recruited for the experiment, of which 10 had to be excluded due to technical difficulties with the robot. Thus the data of 105 participants were used for analysis (*mean age* = 22.8, *SD* = 3.17) with ($N = 21$) participants per condition. All participants were recruited by the INSEAD-Sorbonne University Behavioural Lab under ethics approval

Questions <i>Used, French</i>	Questions <i>English</i>
Pepper est gentil	Pepper is friendly
Pepper est chaleureux	Pepper is warm
Pepper est aimable	Pepper is likeable
Pepper est accessible	Pepper is approachable
Je demanderais volontiers des conseils à Pepper	I would ask Pepper for advice
J'aimerais avoir Pepper comme collègue	I would like Pepper as a colleague
J'aimerais avoir Pepper comme colocataire	I would like Pepper as a housemate
J'aimerais que Pepper et moi soyons amis	I would like to be friends with Pepper
Pepper et moi sommes similaires	Pepper is similar to me

Table 5.2 Likeability Scale ranked from 1 to 7

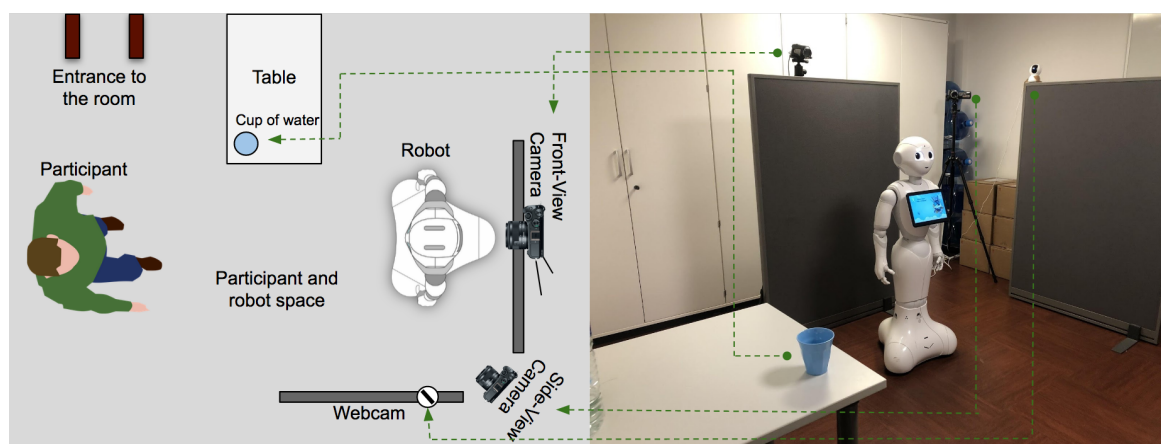


Fig. 5.6 Schematics of the experimental room set-up with the robot during the travel agent scenario

by INSEAD Institutional Review Board. All participants were native French speakers and signed a consent form to participate. Separate consent was obtained for the use of video data. The entire experiment took about 20 minutes to complete, including filling the questionnaires at the end. As a compensation for their time each participant received 6 euros.

Participants were randomly assigned to one of the five conditions (21 per condition). In addition, the participants were assigned randomly across the hours of the day to make sure half of the users in each condition interacted in the morning before lunch time while the other half in the afternoon after lunch.

Upon arrival, the participants filled out the consent forms. Then, one of the experimenters introduced the experiment by explaining that they will interact with a robot that will help them plan their vacation as if it were their real holiday. They were informed the interaction would last 5-7 minutes and then they would have to fill a questionnaire, which included the

liking scale as well as scales not relevant to the current study. They were also advised to speak loudly and articulate clearly in order to avoid any audio or voice recognition problems. Participants were then led to a room with the robot, as seen in Figure 6.2, and were asked to place their belongings on the side and stand wherever they wished. Shortly after the experimenter leaves the room, the interaction began. Besides the front and side cameras set-up in the room, there was a webcam streaming the interaction live. The robot as shown previously is completely autonomous, the videos and webcam are for recording and monitoring. Once the interaction was over, the participants were asked to fill the questionnaire in a different room and then they were debriefed on the study and were given their participation compensation.

5.3.4 Hypotheses

To evaluate the modalities in a multi-modal interaction and their effects on perceived social intelligence, the following hypotheses were formulated. First, the proxemics implemented respected the personal distance established by [65] and visibility and safety [99] while still initiating the interaction by approaching the user within his/her gaze zone [151]. Following this, H1 was suggested:

- H1: Social distances established by the robot would be maintained throughout the interaction in all conditions except *Minus Proxemics*.

Second, social gaze aversion was shown to play a major role in *intimacy* regulation during human-human interactions to elevate the comfort of speakers ([85], [16]). In addition, gaze aversion for turn-taking functioned as a social cue to hand the conversational floor to the user and thus making him/her the speaker. Furthermore, gaze aversion is practiced by humans specially when listening in order minimize the negative perception attributed to staring and to promote the comfort of the speaker ([1], [36]). While Andrist et al. [11] did not find that a social robot with proper timings for gaze aversions increased self-disclosure and comfort in humans more than a social robot with badly timed gaze aversions, we hypothesise that gaze mechanisms supported by multi-modal behaviors would elicit more self-disclosure from the participants, such that:

- H2: Time participants spend speaking in the self-disclosure segment would be the shortest in the *Minus Gaze* condition relative to the other conditions.

Third, gesture and joint attention through gaze have shown to be modalities used to communicate and point at an object of reference in an interaction as well as asking to grab the object referred to ([22], [10] [157]). With H3 formulated as:

- H3: Water suggestions are more likely to be taken when participants interact with a robot performing *social gaze mechanisms* and *gestures*.

Fourth, gestural alignment was proposed to measure the extent to which an interaction shapes the behavior of the user and the smoothness of the interaction [24, 35]. As such, the following hypotheses were formulated looking into gestural alignment at the greeting and termination phases of the interaction in order to understand the possible change in gestural alignment behavior of the users. The use of back-channels, which include nodding and verbal content, throughout the interaction were also analysed as part of gestural alignment:

- H4a: Gestural alignment in the greeting and termination phases would be least present in the *Minus Gesture* condition.
- H4b: *the complete multi-modal behaviour* condition would have the most participants who at the beginning did not greet the robot but at the end did close the interaction with the robot whether verbally or non-verbally.
- H4c: Back-channeling throughout the interaction would be least performed by participants in the *Minus Gesture* condition.

Fifth, all the modalities combined make up multi-modal social cues designed to facilitate a more natural and friendly interaction. It was hypothesized that that would additionally have an effect on the subjective attitude of the users.

- H5: The *condition with all modalities and social dialogue* would score higher on the likeability scale questionnaire.

5.4 Results

5.4.1 Distances from Pepper

First, looking into the distances established throughout the interaction, one would expect no difference in the initial distance, as no interaction has yet occurred, but rather a difference after the navigation of the robot. We therefore conducted Kruskal-Wallis H-tests for the average distance maintained from the robot at the beginning of the interaction (prior to the social navigation phase), shown in Figure 5.7, as well as the average distance maintained after the social navigation phase until the end of the interaction, shown in Figure 5.8. The former model yielded no significant difference for condition, however, the latter revealed a significant effect of condition after the social navigation phase, see Table 5.3.

Follow up pairwise comparisons with Dwass-Steel-Critchlow-Flinger correction revealed participants in the minus proxemics condition stood significantly further away from the robot

throughout the interaction than in all other conditions (all $W's \geq 4.68$, all $p's < 0.008$). No other differences between conditions were significant. This denotes that participants maintained the close social distance due to the proxemics established by the robot, thus validating **H1**.

Table 5.3 Kruskal-Wallis Tests for the Effect of Condition on Initial and Maintained Distances

Outcome	χ^2	df	p	ϵ^2
Initial Distance	2.40	4	.663	0.02
Distance Maintained	23.92	4	<.001	0.23

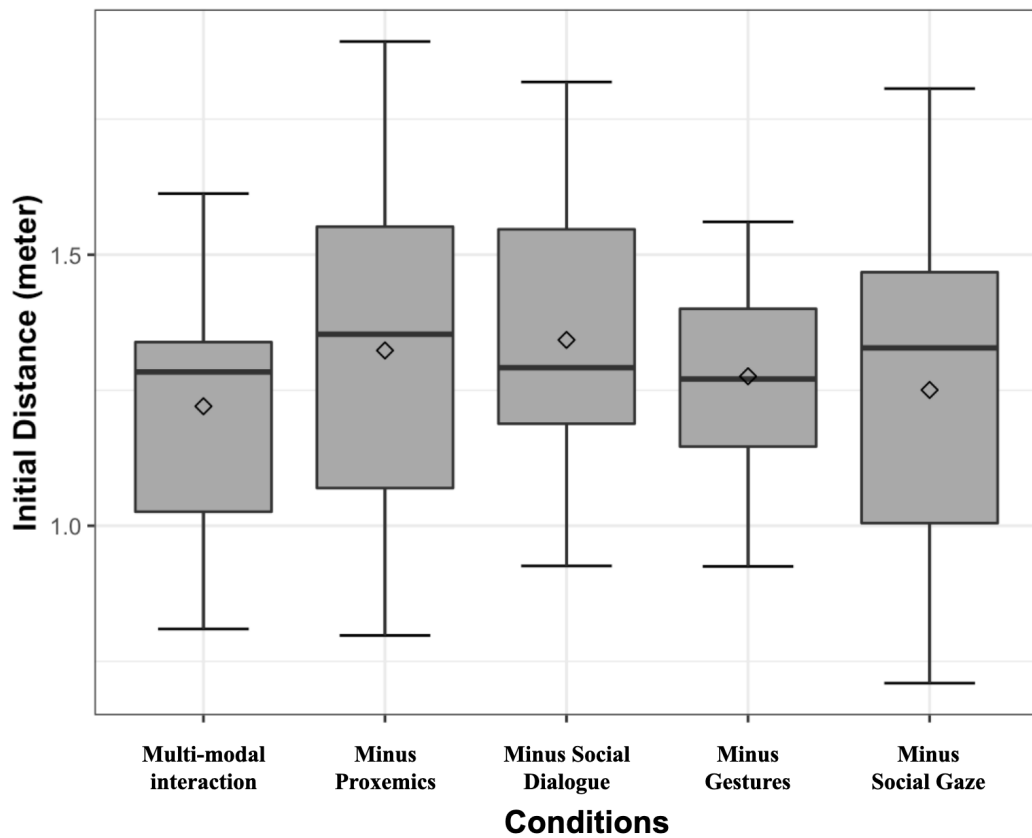


Fig. 5.7 Initial Distance (meters) by participants per condition at the beginning of the interaction with the robot before proxemics

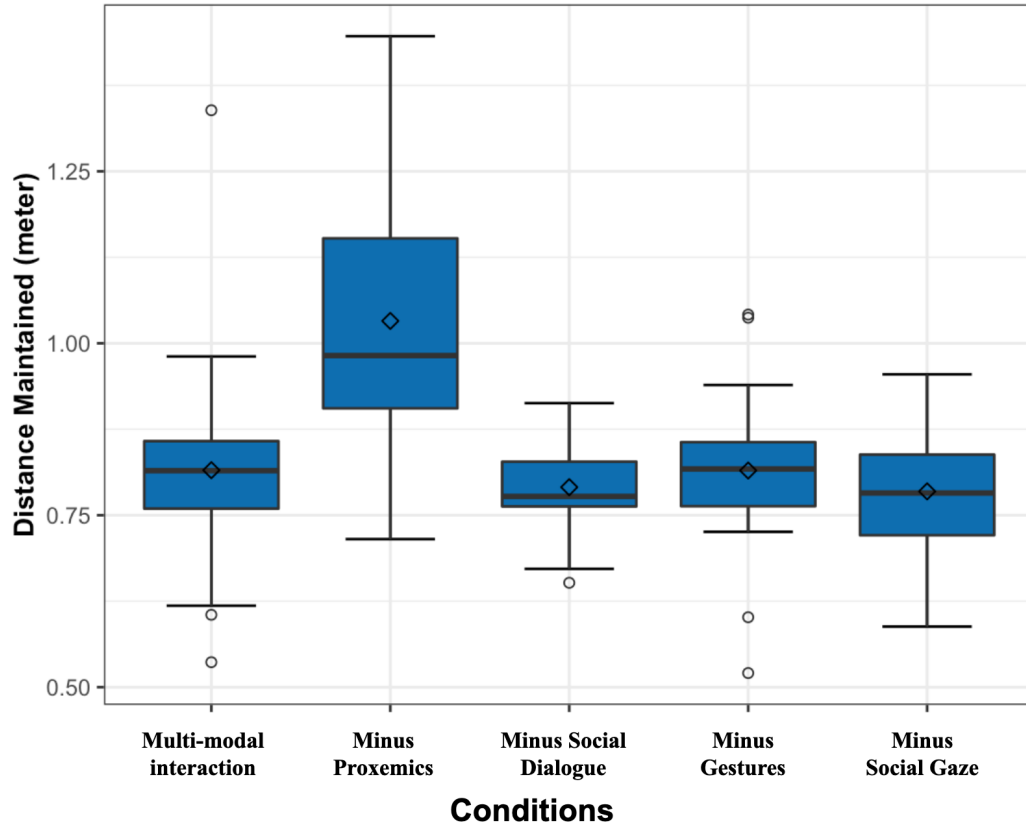


Fig. 5.8 Maintained Distance (meters) by participants per condition throughout the interaction with the robot after proxemics.

5.4.2 Self-Disclosure

Second, we examined the effect of modalities on self-disclosure by looking into how comfortable a human was in sharing information about themselves when asked by the robot. We again conducted a Kruskal-Wallis H-test to evaluate the total speaking time (in seconds) the user spent answering the robot's open-ended question to talk about themselves, shown in Figure 5.9. The average total speaking time was $M = 11.55$, ($SD = 6.05$) seconds. However, there was no significant effect of modality, $\chi^2(4) = 4.09, p = .394, \epsilon^2 = 0.04$. As such, **H2** was not supported. The number of pieces of information given, which is the number of new facts or opinions revealed and provided by the user about him/herself, was annotated and measured for each condition. The following is an example of how the data was annotated: if a participant after the open-ended self-disclosure question answered "I like hanging out with my friends...I like watching movies", then this was annotated as two pieces of information since two facts and/or opinions were revealed about the participant. The average number of

pieces of information given was $M = 3.429$, ($SD = 1.6$), but there was again no significant effect of condition, $\chi^2(4) = 1.76$, $p = .780$, $\epsilon^2 = 0.02$.

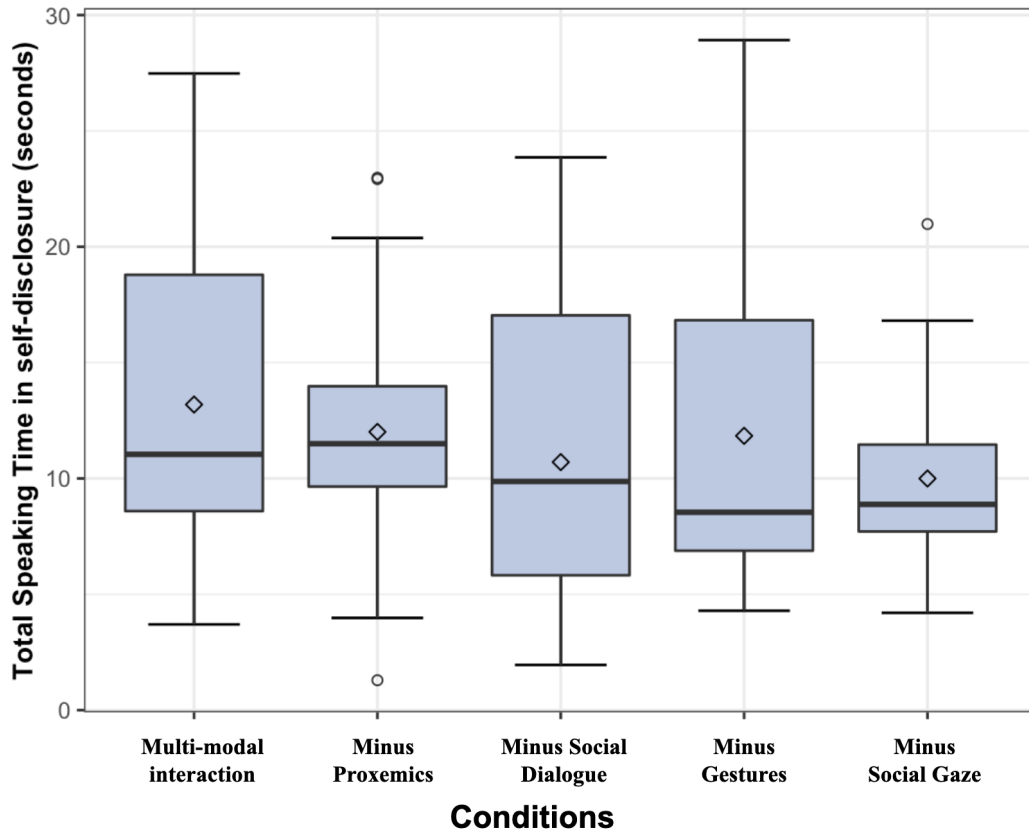


Fig. 5.9 Total speaking time (seconds) of participants per condition during self-disclosure open-ended question

5.4.3 Accepting Water Offered

The number of participants who accepted the water offered in each condition is shown in Figure 5.10. It is important to note that *absent* in this figure and those that follow refer to the number of times the observed behavior was absent in each condition. For instance, in Figure 5.10, the grey or *absent* plots indicate the number times the water suggestions was *not* taken (and as such absent). We ran a binomial logistic regression with condition as the predictor variable, and accepting the water offered as an outcome variable. The overall test of condition was marginally significant, $\chi^2(4) = 8.18$, $p = .09$, McFadden's Pseudo $R^2 = 0.06$. Follow up pairwise comparisons with Tukey's correction revealed a marginally significant difference between the complete multimodal condition and the minus proxemics condition,

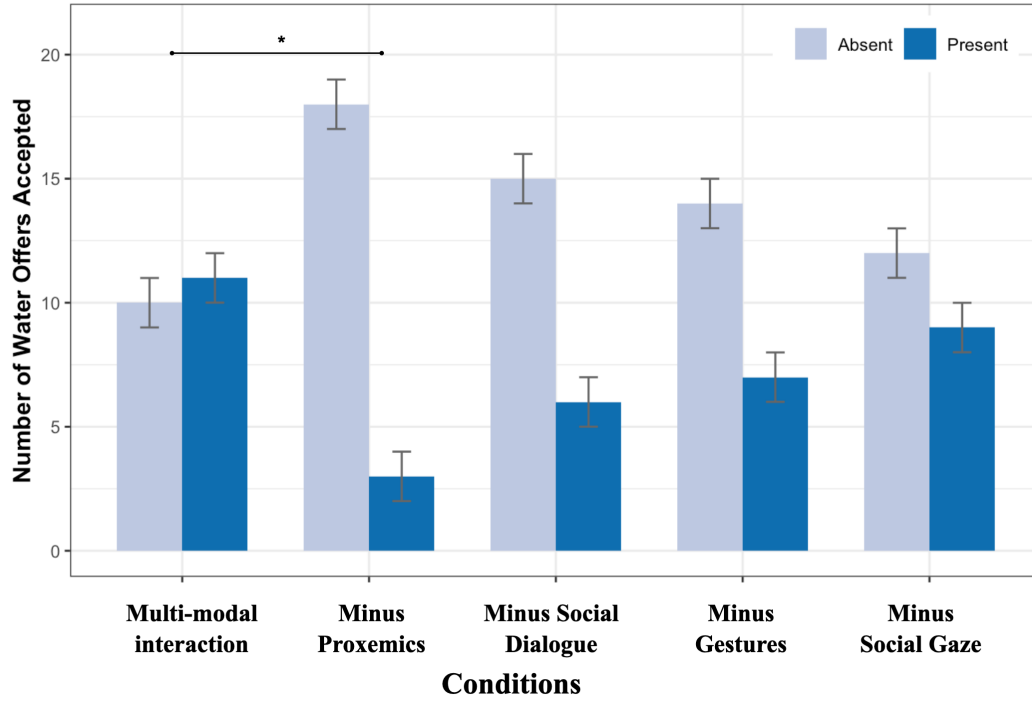


Fig. 5.10 Number of times the water offer was accepted by participants per condition. *Present* refers to appearance of the participant's behavior of accepting the water offered by the robot by grabbing the cup and/or drinking the water, while *Absent* refers to lack of this behavior and as such not accepting the water offered by the robot.

$\beta = 1.89, p = .095$. The odds of accepting the water offered in the minus proxemics condition were 0.15 [0.03, 0.62] times less than in the complete multi-modal condition.

5.4.4 Social Behavior

We again constructed binomial logistic regression models assessing the effect of condition on opening and closing waves. For both opening and closing waves, the logistic regression model was significant for condition, $\chi^2(4) = 19.20, p < .001$, McFadden's Pseudo $R^2 = 0.13$ and $\chi^2(4) = 35.54, p < .001$, McFadden's Pseudo $R^2 = 0.25$, respectively. We conducted follow up pairwise comparisons with Tukey's correction. As the closing wave model exhibited complete separation (i.e., no participants in the minus gesture condition waved goodbye), we further applied Firth's bias reduction method for this model. For both opening and closing waves, participants were significantly less likely to wave in the minus gesture condition than in the complete multi-modal condition and the minus Social Dialogue condition (see Table 5.4). No other comparisons were significant. See Figures 5.11 and 5.12.

Beyond waving gestures in greeting and closing the interaction, Table 5.5 shows the presence and absence of all greeting and/or closing turns made by users, whether verbally or non-

verbally. We classified participants behaviour as either consistent-social (both greeting and closing turn). consistent-nonsocial (neither greeting nor closing), inconsistent-social (no greeting turn, but a closing turn) or inconsistent-nonsocial (greeting turn, but no closing turn). However, a chi-square test comparing participants consistency in social behaviour did not reveal any differences between conditions, $\chi^2(16) = 13.11, p = .664$, Kramer's $V = 0.194$ [0.00, 0.33].

In addition, to assess the effect of modalities on behavioral alignment, we analysed the back-channeling performed by the participants in each condition. A binomial logistic regression was conducted to analyze the effect of the modalities on the number of participants who performed back-channeling, as shown in Figure 5.13. The regression model was statistically significant with $\chi^2(4) = 12.90, p = .012$, McFadden's Pseudo $R^2 = 0.09$. Evaluation of the log odds with Tukey's correction revealed participants in the minus gaze condition were less likely to produce back-channels than in the complete multi-modal condition, see Table 5.6.

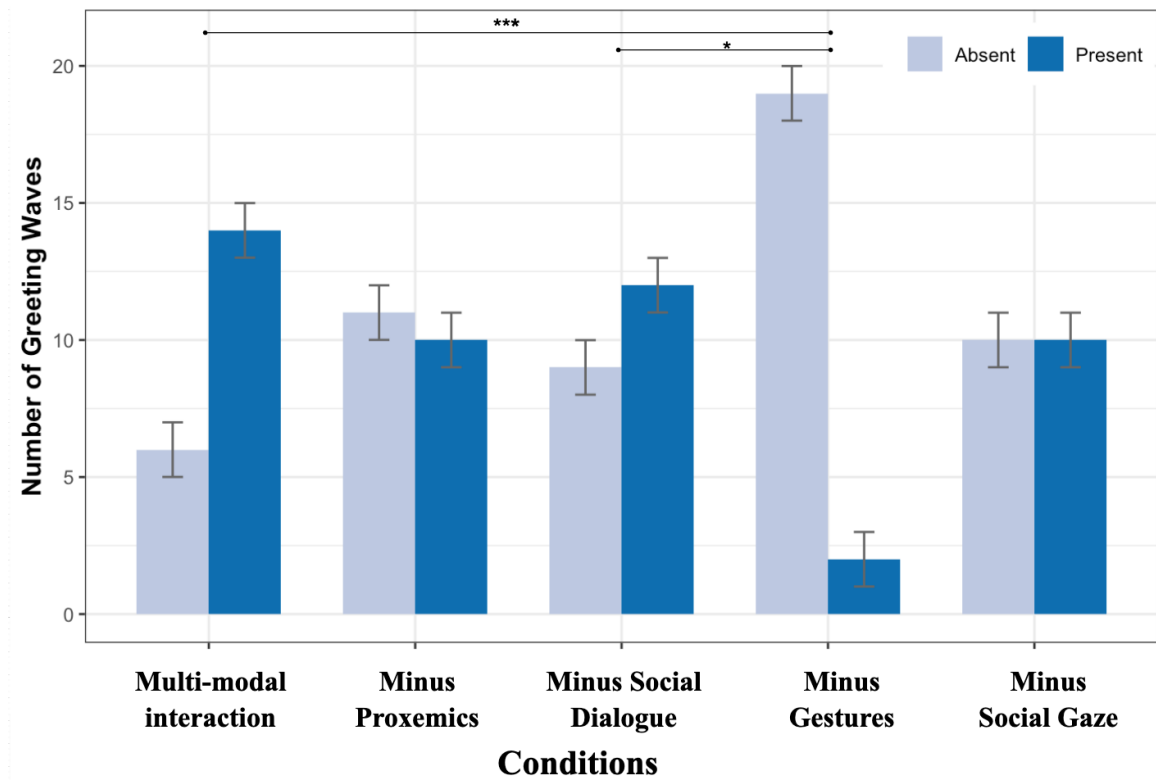


Fig. 5.11 Number of *greeting* waves performed by participants at the beginning of the interaction (with 95% CI errors). *Present* refers to appearance of the participant's behavior of waving to the robot at the beginning of the interaction, while *Absent* refers to lack of this behavior, e.g., the participants not waving to the robot.[*] significant at the $p < .05$ level. [***] significant at the $p < .001$ level

Table 5.4 Comparison of the likelihood of producing opening and closing waves in each condition

Contrast	Log Odds (SE)	z	p -value
Opening Wave			
Minus Gesture vs Multi-Modal Interaction	-3.10 (0.89)	-3.49	.005**
Minus Gesture vs Minus Gaze	-2.25 (0.87)	-2.60	.071
Minus Gesture vs Minus Social Dialogue	-2.54 (0.86)	-2.94	.027*
Minus Gesture vs Minus Proxemics	-2.16 (0.86)	-2.50	.090
Multi-Modal Interaction vs Minus Gaze	0.85 (0.66)	1.28	.704
Multi-Modal Interaction vs Minus Social Dialogue	0.56 (0.66)	0.85	.915
Multi-Modal Interaction vs Minus Proxemics	0.94 (0.66)	1.44	.602
Minus Gaze vs Minus Social Dialogue	-0.29 (0.63)	-0.46	.991
Minus Gaze vs Minus Proxemics	0.10 (0.63)	0.15	.999
Minus Social Dialogue vs Minus Proxemics	0.38 (0.62)	0.62	.972
Closing Wave †			
Minus Gesture vs Multi-Modal Interaction	-4.35 (1.54)	-2.83	.038*
Minus Gesture vs Minus Gaze	-3.95 (1.53)	-2.58	.074
Minus Gesture vs Minus Social Dialogue	-4.63 (1.54)	-3.01	.022*
Minus Gesture vs Minus Proxemics	-3.30 (1.53)	-2.15	.197
Multi-Modal Interaction vs Minus Gaze	0.40 (0.65)	0.61	.973
Multi-Modal Interaction vs Minus Social Dialogue	-0.28 (0.67)	-0.42	.994
Multi-Modal Interaction vs Minus Proxemics	1.05 (0.65)	1.62	.482
Minus Gaze vs Minus Social Dialogue	-0.68 (0.66)	-1.03	.840
Minus Gaze vs Minus Proxemics	0.65 (0.64)	1.03	.842
Minus Social Dialogue vs Minus Proxemics	1.33 (0.66)	2.03	.251

† With Firth's bias reduction method

* significant at the $p < .05$ level** significant at the $p < .01$ level

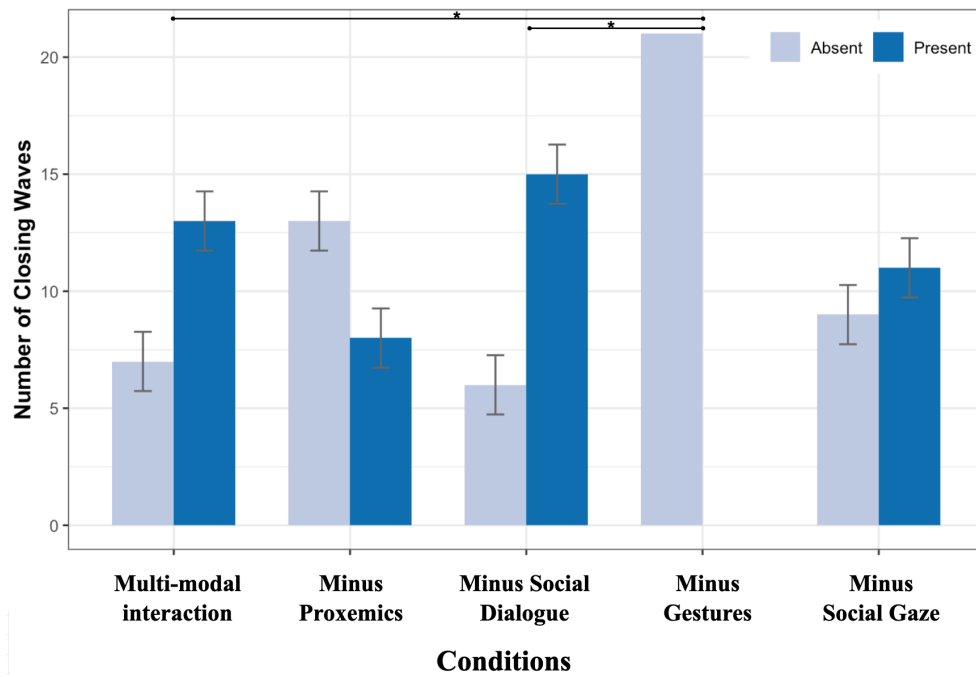


Fig. 5.12 Number of *closing* waves performed by participants at the end of the interaction (with 95% CI errors). *Present* refers to appearance of the participant's behavior of waving to the robot at the end of the interaction, while *Absent* refers to lack of this behavior, e.g., the participants not waving to the robot. [*] significant at the $p < .05$ level

Status	Multi-modal	Minus	Minus	Minus	Minus	Total
of Greeting/Closing:	Interaction	Proxemics	Social Dialogue	Gestures	Gaze (n_5)	(N)
	(n_1)	(n_2)	(n_3)	(n_4)		
Greeted robot & Closed Interaction	10	15	16	13	14	68
Greeted robot but Did not Close Interaction	4	2	1	1	2	10
Did not Greet robot but Closed Interaction	5	3	1	5	2	16
Did not Greet robot & Did not Close Interaction	1	1	3	2	1	8

Table 5.5 Number of greeting and/or closing turns, which may be verbal or non-verbal, present and absent in each modality in beginning and end of the interaction. The rows represent the statuses in each condition as follows: 1) participants who greeted the robot and also closed the interaction (no significance) 2) those who greeted the robot but did not close the interaction 3) those who did not greet the robot but closed the interaction 4) those who did not greet the robot nor did they close the interaction

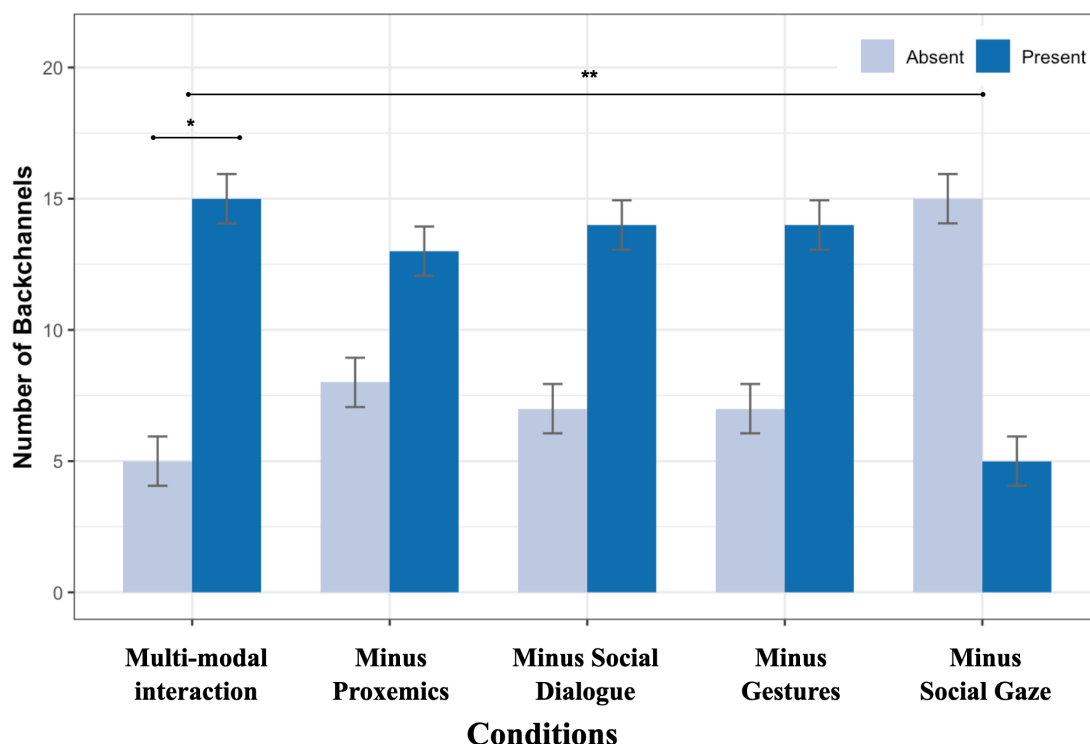


Fig. 5.13 Number of back-channels performed while interacting with the robot in each condition (with 95% CI errors). *Present* refers to appearance of the back-channels performed by the participants, while *Absent* refers to lack of back-channels detected during the interaction.[*] significant at the $p < .05$ level. [**] significant at the $p < .01$ level

Table 5.6 Comparison of the likelihood of producing back-channels in each condition

Contrast	Log Odds (SE)	z	p -value
Minus Gesture vs Multi-Modal Interaction	-0.41 (0.69)	-0.59	.977
Minus Gesture vs Minus Gaze	1.79 (0.69)	2.58	.073
Minus Gesture vs Minus Social Dialogue	0.00 (0.66)	0.00	1.00
Minus Gesture vs Minus Proxemics	0.21 (0.65)	0.32	.998
Multi-Modal Interaction vs Minus Gaze	2.20 (0.73)	3.01	.022*
Multi-Modal Interaction vs Minus Social Dialogue	0.41 (0.69)	0.59	0.98
Multi-Modal Interaction vs Minus Proxemics	0.61 (0.69)	0.80	.899
Minus Gaze vs Minus Social Dialogue	-1.79 (0.69)	-2.58	.073
Minus Gaze vs Minus Proxemics	-1.58 (0.69)	-2.31	.140
Minus Social Dialogue vs Minus Proxemics	0.21 (0.65)	0.32	.998

* significant at the $p < .05$ level

5.4.5 Liking

A Kruskal-Wallis H-test for participants subjective evaluations of their liking of the robot revealed no differences between any of the conditions, shown in Figure 5.14, $\chi^2(4) = 3.84, p = .428, \epsilon^2 = 0.04$; failing to support **H5**. Further exploratory investigation was done to look into behavioral outcomes that might represent liking. First, Figure 5.15 shows the number of addressee terms used to address the robot in each condition. In the french language, the pronoun “tu” (referring to “you”) is used in informal and/or friendly contexts, whereas, the pronoun “vous” is used for formal and/or acquaintance contexts. See Table 5.7 for the frequency of each mode of address used by participants. A chi square test comparing participants mode of address towards Pepper in each condition was significant $\chi^2(12) = 23.01, p = .028$, Kramer’s $V = 0.27, [0.00, 0.32]$. Follow up tests with FDR correction, however, did not reveal specific differences between conditions. As this analysis was exploratory, we then relaxed the need for correction with multiple comparisons. Without correction, there was a significant difference between the minus Social Dialogue and minus proxemics conditions.

Second, Figure 5.16 shows in each condition the number of participants who made utterances while using the tablet (despite it being clear that speech was not needed to carry out choice selection on the tablet of the robot). For instance, listening mode and speech recognition in the robot was activated when the robot asked open ended questions; whereas for making choices regarding the planning of the destination in this scenario, the interface to select the preference was by using the tablet (as was instructed by the robot in the beginning). There was no significant difference in how likely participants were to talk to Pepper in addition to using the tablet, $\chi^2 = 5.18, p = .270$, McFadden’s Pseudo $R^2 = 0.04$.

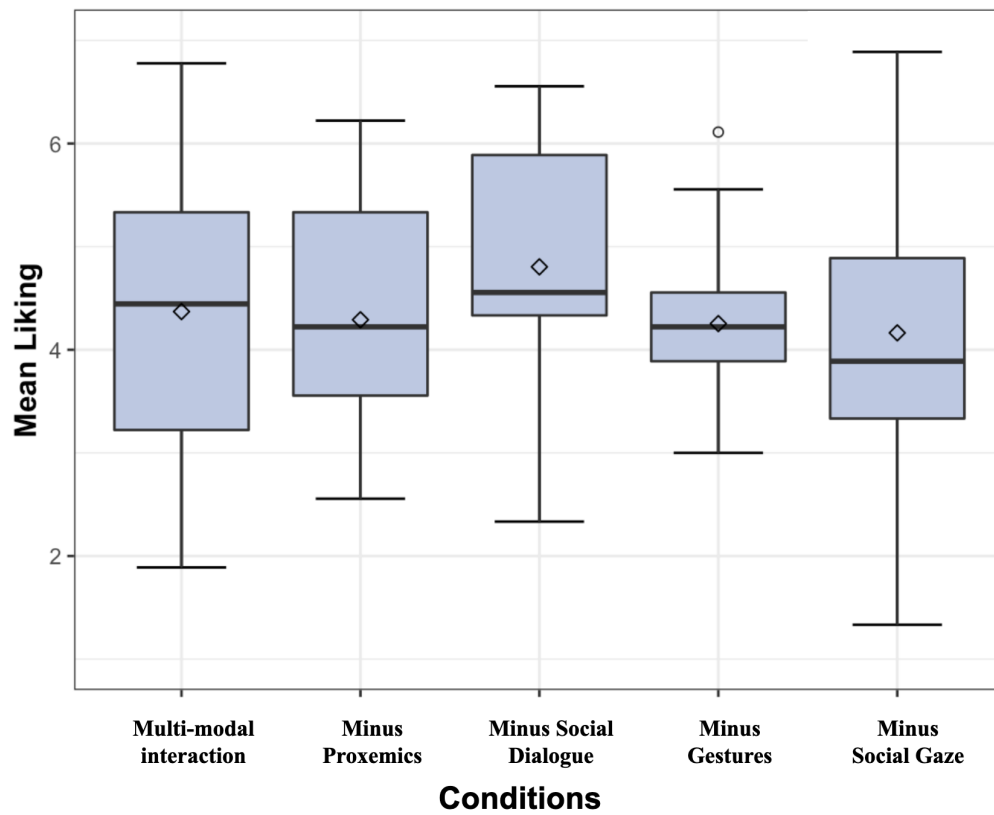


Fig. 5.14 Mean Liking ranked by participants for the robot in each condition

Table 5.7 Frequency of modes of address used by participants towards Pepper in each condition

Condition	Mode of Address			
	Tu	Vous	Pepper	Nothing
Multi-Modal Interaction	1	1	0	19
Minus Gaze	1	3	0	17
Minus Gesture	2	3	0	16
Minus Social Dialogue	0	1	3	17
Minus Proxemics	5	1	0	15
Totals	9	9	3	83

Table 5.8 Comparison between modes of address used by participants towards Pepper in each condition

Contrast	Raw <i>p</i> -value	FDR- corrected <i>p</i> -value
Multi-Modal Interaction vs Minus Gaze	.798	.886
Multi-Modal Interaction vs Minus Gesture	.593	.740
Multi-Modal Interaction vs Minus Social Dialogue	.232	.386
Multi-Modal Interaction vs Minus Proxemics	.229	.386
Minus Gaze vs Minus Gesture	1.00	1.00
Minus Gaze vs Minus Social Dialogue	.223	.386
Minus Gaze vs Minus Proxemics	.207	.386
Minus Gesture vs Minus Social Dialogue	.122	.386
Minus Gesture vs Minus Proxemics	.347	.496
Minus Social Dialogue vs Minus Proxemics	.022*	.222

* significant at the $p < .05$ level

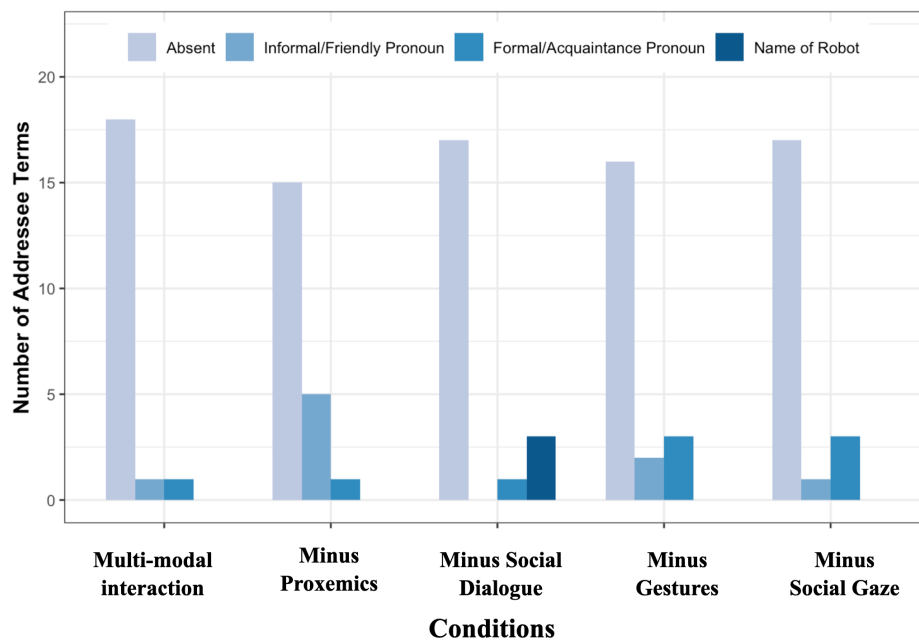


Fig. 5.15 Number of addressee terms used to address the robot in each condition. The addressee terms are placed in four categories: absent, no term was used to directly address the robot, informal/friendly pronoun, formal/acquaintance pronoun, and name of the robot

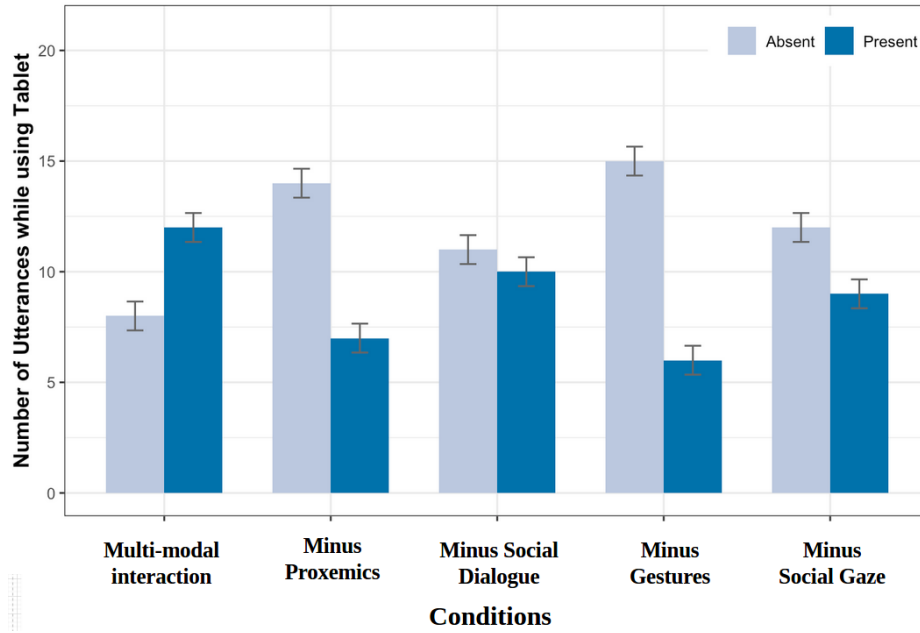


Fig. 5.16 Number of Utterances while using the tablet of the robot in each condition. *Present* refers to detection of utterances/speech while tablet was used by the participants, while *Absent* refers to lack of detection.

5.4.6 Voice Recognition Errors

Although we attempted to limit the amount of autonomous voice recognition, the introductory phase included some reciprocal interaction between Pepper and the participant (e.g., asking “how are you”). Thus, there was still some potential for voice recognition errors to occur. Although not initially part of the experimental design, based on observations of instances where voice recognition errors occurred we decided to explore if these (naturally occurring) errors had any effect on participants behaviour.

A one-way Kruskal Wallis H-test performed on the number of voice recognition errors occurring per condition was non-significant, $\chi^2(4) = 2.04$, $p = .727$, $\eta^2 = 0.02$, indicating there was no difference in the number of voice recognition errors occurring between conditions.

Significant negative correlations were identified between the number of voice recognition errors, the total time the participant spent talking during the self-disclosure phase $\rho = -0.20$, $p < .05$ and the number of pieces of information they disclosed $\rho = -0.28$, $p < .01$.

5.5 Discussion

First, *H1* was supported, showing the influence of proxemics on the distances maintained by the users throughout the interaction. Prior to the navigation of the robot, participants chose

to stand far from the robot at a ($M = 1.28, SD = 0.26$) meters distance, with no significant difference between conditions. The initial distance chosen was within the social distance defined by Hall [65] and did not give accessibility to the robot's tablet. In the conditions where the robot navigated to establish the personal distance of 0.85 meters, distances maintained during the rest of the interaction until the end were much closer. Conversely, in the *minus proxemics* condition, participants kept a further distance from the robot. Proxemics once again played an influential role on the behavioral outcomes of the interaction and was the main reason users kept a close distance to the robot.

Second, $H2$ was not supported, *social gaze* mechanisms did not elicit an increase in self-disclosure speaking time nor the amount of information the user revealed about themselves. Further investigation was held to interpret what might have affected the self-disclosure speaking turn of the users. It was found that *voice recognition errors* significantly predicted total speaking time of participants and the amount of information shared. While gaze aversions and their respective functions play a guiding role in intimacy regulation and comfort in self-disclosure in human-human interactions, findings in this study seem to show that for human-robot interactions technological voice recognition errors precede gaze aversions in governing behavioral outcomes for such contexts. This shows that getting the robot technologically ready may have a great impact on how naturally a user answers an open ended question about themselves rather than how close a robot's subtle behavior is to a human.

Third, while $H3$ was not fully supported, the results gave an insight into the effect of *proxemics* modality. The condition in which there was least water suggestions taken was in the *minus proxemics* condition. Even though, the state of the art has been focused on using deictic gesture and joint attention gaze for pointing at objects to grab for task-oriented scenarios ([10], [157]), there was no significant difference for these modalities in this study. In addition, the suggestion of object grabbing in this chapter was more focused on its social context and implications. It was shown that the participants not only took the object suggested by robot, in this case the water cup, but also drank the water. It may perhaps be linked to the perception of the user to the robot's *situational awareness*, which is the ability to perceive and infer knowledge from the surrounding environment [21]. There is a need for future work to better understand the potential of proxemics on object manipulation in the shared environment between the user and robot.

Fourth, while $H4a$ was validated, **H4b** was not validated and $H4c$ was partially supported. The lack of social gestures significantly affected the behavioral alignment for the greeting and closing of the interaction in the *minus gestures* condition. In the greeting part of the interaction, the *minus gestures* condition had significantly less wave gestures performed to greet the robot than in the complete multi-modal and minus social dialogue conditions. For

closing waves this difference was even more extreme, with, no wave gestures performed in closing the interaction with the robot in the minus gestures condition. Even further, In Table 5.5, while the number of participants that did not greet the robot but eventually performed a closing turn at the end of the interaction were highest in the *multi-modal interaction* and *minus gestures* conditions, there was no significant difference. This may also imply that even though there was no behavioral gesture done in the closing of the interaction in the *minus gestures* condition, there was a verbal closing turn.

On the other hand, $H4c$ was partially supported. While it was not the *minus gestures* condition that had the least amount of back-channeling alignment performed by users as hypothesized, it was instead the *minus social gaze* condition. This may indicate that gesture mirroring was not the main cause of back-channeling alignment, but rather how naturally the interaction flowed. Conditions with social gaze mechanisms included turn-taking and floor-holding which hold cognitive functions and were accompanied by very short pauses in speech. The users may have performed more back-channeling during these conditions as it was a natural human behavior and as a way to provide the robot feedback that they were in fact still listening to its speech and aligned in the interaction. Thus, the social gaze plays a role in shaping the human-robot interaction seem more instinctive to the human and in forming alignment.

Fifth, a self-reported questionnaire was used to measure liking or ‘likability’ of the robot and it was hypothesized that the *multi-modal interaction* condition would score higher; however $H5$ was not supported. Further behavioral outcomes were annotated and analyzed that might be related to liking of the robot. First, the way the participants addressed the robot was studied. The experiment took place in French with native french speakers and in the french dialect the “you” pronoun is represented by “vous” for formal set-ups and/or with acquaintances and by “tu” for rather informal set-ups and/or with friends. The minus social dialogue condition was significantly different to the minus proxemics condition. The *minus social dialogue* condition was the only one to have users address the robot by its name, e.g., here being “Pepper”. In addition, the *minus proxemics* condition had the highest number of participants using informal/friendly pronouns. Further research needs to be done to better comprehend what that would signify but it can be concluded at this point that modalities affect the terms participants exercise in the interaction with the robot. Second, during the interaction, the robot is in listening mode at only two phases: at the beginning during the social small talk shown in Table 5.1 and at the open ended question to measure self-disclosure. At all other times during the interaction, the robot was not in a listening mode and its tablet was required to be used by the participant to answer the questions asked by the robot. However, it was noted that some participants chose to talk while using the tablet and to sometimes justify their choices to the robot and discuss their thought process

out loud. While no significant difference was found in the results, the *multi-modal interaction* condition had the highest number of participants who chose to also talk while using the tablet and the *minus gestures* condition had the least. This may give an insight into the effect of *social gestures* on how social participants were with the robot. These findings hint into the type of relationships users formed with the robot based on the multi-modal behaviors they interacted with. There seems to be more to discover and investigate in future works.

5.6 Conclusion

Non-verbal behavior plays a key role in human communication not only by reinforcing and enhancing speech in diverse formats of the interaction, but also by carrying fundamental functions in communication that can stand-alone from speech. However, this non-verbal behavior is not made of only one modality but rather of multi-modalities all composed together to serve their purpose. For this reason, while studying each modality separately may lead to improving human-robot interaction, a deeper understanding of the different modalities when performed together and their combinations as well as their interaction outcomes is imperative for effective use of the multi-modalities of robots in maximizing targeted outcomes. This chapter presented work attempting to build such an understanding. The process involved implementing a system of multi-modalities including social gaze mechanisms, different types of gestures, proxemics for navigation in initiating conversations, and social dialogue followed by an evaluation study where participants interacted with the robot in a travel agent scenario. The system and methodology presented in this chapter can be as well utilized on other robots. The results showed various insights into the contributions of modalities in a multi-modal interaction onto several notable behavioral outcomes of the users, including taking physical suggestions, distances maintained during the interaction, wave gestures performed in greeting and closing, back-channeling, how the robot is addressed, and how socially it is treated. It can be concluded that certain modalities in multi-modal behaviors particularly influence the outcomes of the interaction, and at times not in the same way as seen in the state-of-the-art of the modality on its own. Notably, this chapter showed how multiple modalities can be combined in an interaction and how subtracting each modality at a time revealed insights about the effect of that modality. For instance, it is now clear how proxemics influence the distances maintained during an interaction and the probability of the user accepting the robot's offer. In addition, social gestures can predict how humans greet the robot and close interactions with it and the utterances the user makes while using other modalities of the robot such as the tablet. Moreover, social gaze shaped how naturally humans back-channel when interacting with the robot but did not have an effect on how much they disclose to the robot. All these findings may lead to further understanding on human-robot interaction and

how multi-modal behavior can be used to increase the perceived social intelligence of the robot.

While Chapter 4 demonstrated how one modality from the human's behavior can be used to adapt another modality in the robot's behavior to interaction changes such as changes in group dynamics, Chapter 5 synthesized and investigated the effect of modalities making up the multi-modal behaviors of the robot on the behavioral outcomes of the participant in a human-robot interaction. Building on the findings of the two previous chapters and aiming at closing the loop, the next Chapter proposed a reinforcement learning model for adaptation and personalization of multi-modal behaviors of the robot based on the human's social signals.

Chapter 6

Adaptation of Multi-Modal Behaviors using Reinforcement Learning

This chapter contains:

6.1	Introduction	86
6.2	RL Model	87
6.2.1	State-Space	87
6.2.2	Action-Space	88
6.2.3	Reward Function	89
6.3	Evaluation	90
6.3.1	Experiment Study	90
6.3.2	Simulation of Human and Interaction	91
6.4	Implementation of Reinforcement Learning Models	94
6.4.1	Multi-Armed bandit	95
6.4.2	Q-Learning	95
6.5	Results	96
6.5.1	Multi-Armed Bandit	96
6.5.2	Q-Learning	103
6.6	Discussion	107
6.7	Limitations	110
6.8	Conclusion	110

6.1 Introduction

Human behavior is fundamentally multi-modal and includes gaze, hand gestures, facial expressions, and body posture among other modalities to express engagement and regulate social interactions [49]. In Chapter 4, it was shown how looking at distances and orientations of individuals in small groups of three forming around the robot allowed the robot to estimate footing, the conversational roles of the participants. Moreover, robots are embodied agents that can interact and manipulate the environment they are in through their multi-modal abilities as shown in Chapter 5. Consequently, multi-modal social signals of humans are rich in information and can enable the robot to learn from them and adapt to changes in interactions [141] and personalize behaviors for different human personalities and cultures to optimize its social intelligence. In parallel, reinforcement learning (RL) is a computational approach to learning from interactions. Thus, in recent years, there has been a growing interest of using RL for adaptation in HRI. From literature review for RL in Social Robotics in Chapter 3, several studies and models are either task dependent or require too large state-spaces to take into account the social signals.

Furthermore, for human-human interactions, adaptation is crucial for interpersonal relationships [33] and the social intelligence of an individual [57]. For instance, people adapt their interaction patterns, behaviors, and conversational styles to one another's to achieve smooth and successful communication and increase the rapport between those participating in the interaction [55]. Similarly, human users expect the agents, robots, and technologies they are interacting with to also adapt to them [108, 19]. In addition, adaptation has shown to improve the usability of these technologies [175]. In human-agent interaction, an adaptive system does not necessarily have the agent learn new behaviors, but rather decide *what* to adapt, by for instance combining different behaviors, and *when* or *when not* to make these modifications [170]. Moreover, in human-computer interaction, adaptation relies on *estimation of utility*, where utility refers to the effectiveness of an adaptation to the user, or how it is assessed to improve the interaction specially when possible costs are considered [170]. While utility may be hard to quantify, RL allows the agent to learn by interacting with the human user and adapt its behavior by investigating the effect of the various multi-modal combinations.

This chapter introduces a RL model that avoids discretizing social signals and large state-space actions by contributing a reward function that integrates the multi-modal social signals from the human, the decisions of the users influenced by the robot, and a cost function on the complexity of performing the multi-modal behavior. In addition, training of RL models for HRI can be very costly and time consuming. Based on the data-set collected

from the study in Chapter 5, this chapter also presents a simulated human in a human-robot interaction for training and evaluation of RL models. Finally, the RL algorithms implemented and studied are Multi-Armed Bandit problems, Upper Confidence Bound, and Q-Learning. The objective is to have a robot that is capable of adapting its multi-modal behavior based on the social signals from the human it is interacting with to create its social influence and intelligence. Figure 6.1 shows an overview of the multi-modal adaptive behavior system using RL.

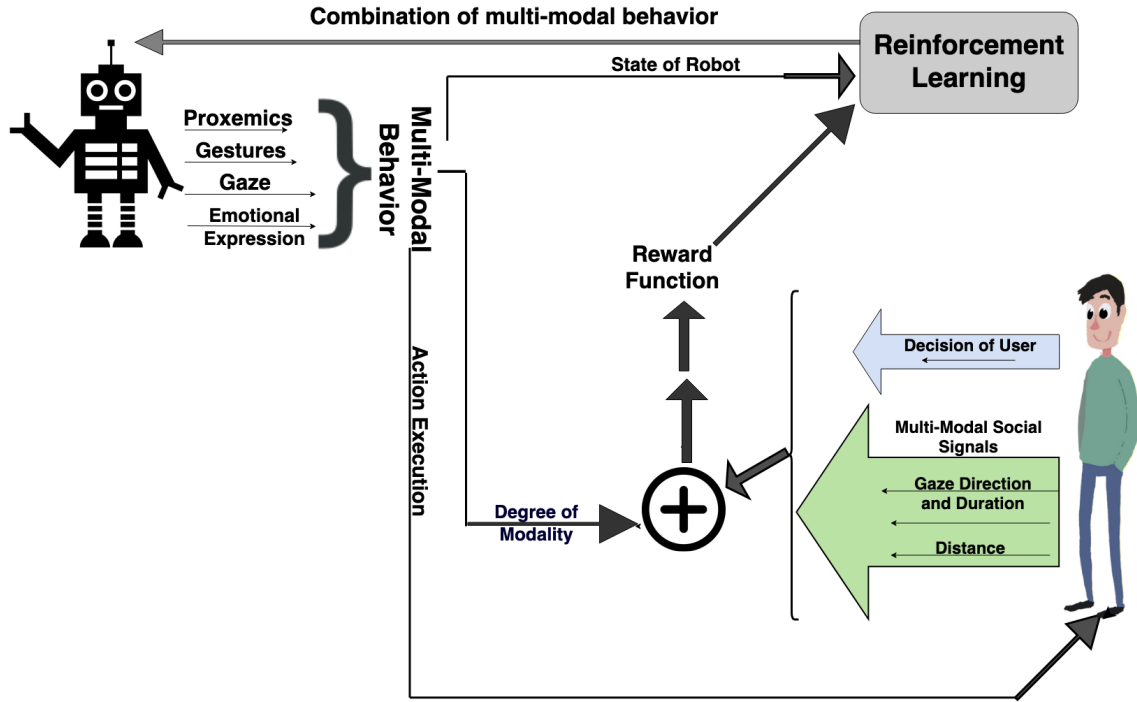


Fig. 6.1 Overview of the multi-modal adaptive behavior system

6.2 RL Model

6.2.1 State-Space

The state-space \mathcal{S} is a one-dimensional nondiscretized vector representing the state of the multi-modal behavior of the robot as shown in Equation 6.1. Thus each state encodes which modalities are currently making up the multi-modal behavior being applied during the behavior generation. \mathcal{S}_{robot} is defined as a four-tuple (*Proxemics*, *Gesture*, *Gaze*, *Emotional Expression*) and each tuple is made up of $\{0, 1\}$, where 0 and 1 represent the deactivation and action of that specific modality respectively. Thus a state s_t is a vector array of length 4

and made of 0's and 1's indicating which modalities are present such that $s_t = (\text{proxemics}, \text{gesture}, \text{gaze}, \text{emotional expression})$. For example, $s_t = (1, 0, 1, 0)$ implies that the behavior is a two-modality behavior made of proxemics and gaze.

$$\mathcal{S} = \mathcal{S}_{robot} = (\text{Proxemics}, \text{Gesture}, \text{Gaze}, \text{EmotionalExpression}) \quad (6.1)$$

Thus the behavior of the robot is a combination of these modalities, where *proxemics* refers to the social navigation of the robot to maintain a close social distance with the user throughout the interaction, *gesture* refers to symbolic, deictic, and beat gestures, *gaze* refers to the functional gaze aversions for turn-taking, floor-holding, turn-yielding, and joint attention, and *emotional expression* refers to the robot showing joyful emotional expressions when the human user takes their recommendation and sad emotional expressions when they do not (inspired by the literature review in Chapter 2). The *proxemics*, *gesture*, *gaze* modalities refer to those designed and implemented in Chapter 5.

6.2.2 Action-Space

To allow the robot to create different combinations of modalities forming its multi-modal behavior, two actions for every modality found in \mathcal{S}_{robot} can be taken: *activate* (+1) and *deactivate* (-1). Thus the action space \mathcal{A} is defined as $A = \{\text{Proxemics}\{-1, +1\}, \text{Gesture}\{-1, +1\}, \text{Gaze}\{-1, +1\}, \text{Emotional Expression}\{-1, +1\}\}$. Thus the action a_t is a vector array. For example, for the robot to display a multi-modal behavior made of proxemics, gaze, and emotion expression starting from a state of no modalities exhibited, the action would be $a_t = (+1, 0, +1, +1)$. However, for simplification purposes, in this study, the agent/robot can only activate or deactivate one modality at each step. The action-space in this study is thus made of 8 possible discrete actions as follows:

1. *proxemics* = +1: add proxemics to current (behavior)
2. *gesture* = +1: add gesture to current state (behavior)
3. *gaze* = +1: add gaze to current state (behavior)
4. *emotional - expression* = +1: add gaze to current state (behavior)

If modality to be added is already present, no change is made

5. *proxemics* = -1: remove proxemics from current (behavior)

6. $gesture = -1$: remove gesture from current state (behavior)
7. $gaze = -1$: remove gaze from current state (behavior)
8. $emotional - expression = -1$: remove gaze from current state (behavior)

If modality to be removed is already absent, no change is made

6.2.3 Reward Function

The reward function takes into account three important considerations. First, it needs to reflect the trade-off between exploring and reaching the optimal multi-modal combination and the degree of modality for each behavior by a cost function for high degrees of modality ($degree_{modality}$). Second, in HRI, social signals of humans change continuously and rarely take the form of explicit feedback. For this reason, the reward function includes the movement distance (d_h), time the human spent looking at an object ($T_{objectlooking}$) referred to by the robot, and time spent the human spent to make a decision ($T_{deciding}$). Third, the reward function also incorporates the decision made by the human (where D can be: +1 for *taking robot's recommendation* and -1 for *did not take robot's recommendation*). The reward function R is shown in Equation 6.2.

$$R = -\kappa(d_h) + \tau \left(\frac{T_{objectlooking}}{T_{deciding}} \right)^{\frac{1}{\phi}} + \beta(D) - \delta(degree_{modality}) \quad (6.2)$$

Social signals of the humans making up the reward function may not be synchronous in time and may not be performed at the same time. For this reason, window frames of 10 seconds, which made be subject to change depending on the robot and use cases, are needed to collect the social signals and calculate the average reward function in this frame.

Furthermore, modalities forming multi-modal behaviors of the robot have an impact on certain behavioral outcomes from the user and their social signals, which are coupled in nature. This played an important role in shaping the *Reward Function* (Equation 6.2) as follows:

$$R = \underbrace{-\kappa(d_h) + \tau \left(\frac{T_{objectlooking}}{T_{deciding}} \right)^{\frac{1}{\phi}}}_{\text{Social Signals}} + \underbrace{\beta(D)}_{\text{Decision Made}} - \underbrace{\delta(degree_{modality})}_{\text{Cost of Modality Degree}}$$

Human
Robot

First, a cost function was included for the distances held by the users as it was shown that proxemics of the human is an important social signal that may influence other behavioral outcomes. As such, the further the participant stands from the robot, the more costly it would be for the robot. Second, while $T_{objectlooking}$, which is the time spent looking at referred object, and $T_{deciding}$, which is the time to make a decision on whether to grab object or take a suggestion are valuable social signals, they are highly coupled and may overlap. For this reason, they were presented in a form of a fraction in the reward function (Equation 6.2). Additionally, the fraction was raised to the power of $(\frac{1}{\phi})$ to smooth the highly fluctuating social signals since human social signals inherently oscillate and this oscillations needs to be decreased to have a clearly reading of the information. The former two make-up the human social signals part of the reward function.

Third, while the the social signals making up the reward function are continuous signals, the decision made by the user of either taking suggestion or the robot or not is a direct discrete feedback given to the robot by the human and as such influencing the reward to be received. If the human takes the recommendation of the robot, then it would positively influence the reward with $D = 1$ but if they do not, then it would cost the robot with $D = -1$. Fourth, a cost function for the degree of modality of the multi-modal behavior of the robot was added such that a behavior including only one modality would have $degree_{modality} = 1$ but a behavior including all four modalities in one multi-modal behavior would have $degree_{modality} = 4$. The former cost function was inserted to create a trade-off between optimizing behaviors for social signals of the user and the cost on the robot on generating high-degree modalities, and to allow the robot to further explore various combinations of modalities.

6.3 Evaluation

This section focuses on discussing the experiment study, the proposed simulation of the human’s behavior in a human-robot interaction based on the robot’s actions, and the implementation of the RL algorithms.

6.3.1 Experiment Study

To train and evaluate the performance, social influence, and social intelligence of the robot, the following experiment was designed. Prior to the experiment, the participants would be informed that they need to gather a list of 10 items for either a travel trip or a survival kit if stuck on an abandoned island or even perhaps a shopping list. Inside the experiment room, the robot would be greeting the human in the middle of the room and on each side of the

room are 10 different objects placed. The aim is to gather the required list while interacting with the robot. At each checkpoint, the robot would recommend one of two items placed on opposite sides of the room. The robot would communicate its recommendation using speech and its nonverbal multi-modal behavior. At the very end of the list of items needed to be gathered, the robot would explain to the paid participant about a certain Non-Governmental Organisation (NGO) working towards a global cause and asks the participants if he/she might be interested in donating its participation fee as a contribution to the NGO. In this study, no physical data collection for the results was done. However, setting up the experiment study was important to better define and implement the simulation of the human-robot interaction and the use case.

6.3.2 Simulation of Human and Interaction

To investigate and tune the weights in the reward function shown in Equation 6.2 as well as to explore simulations as close as possible to real life interactions, the data set used was collected from a previous study, which evaluated the behavioral outcomes and subjective measurements of multi-modal behaviors on a social robot with 105 participants in a seven minute interaction [168], as described in Chapter 5. The former study presented a multi-modal interaction focusing on the following modalities: proxemics for social navigation, gaze mechanisms (for turn-taking floor-holding, turn-yielding and joint attention), gestures (for symbolic, deictic, and beat gestures), and social dialogue.

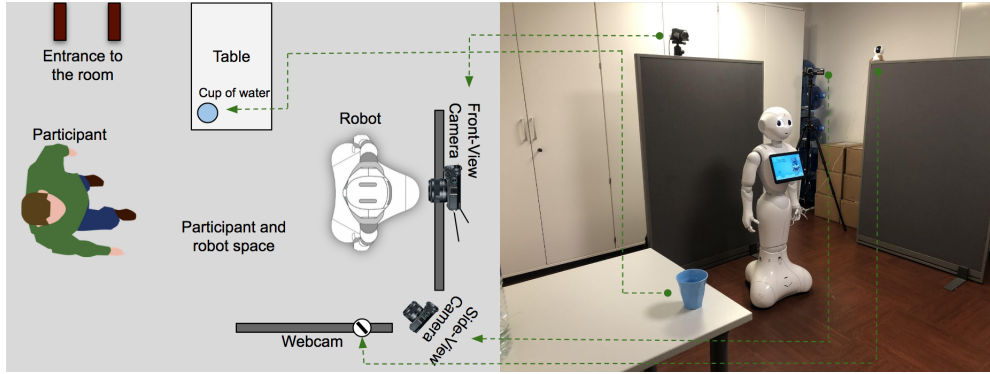


Fig. 6.2 Schematics of the experimental room set-up with the robot of data-set used (Image from Tatarian et al. [168])

Data Collection

The experimental flow of the interaction for the data collection was that explained in Chapter 5. Among the various findings, it was shown that proxemics significantly increased how close participants stood to the robot throughout the entire interaction even if initially they started the interaction further away from the robot. In addition, it was found that proxemics increased the likelihood of participants taking the water offered by the robot. Further investigation of the data collected revealed that participants significantly looked longer at the object referred to (water cup in this case) when they took the water offer than when they did not take the water offer. The data also showed that when gesture or gaze modalities were missing, the decision time demonstrated by users was longer than in conditions when these modalities were present.

Data for Simulation of HRI

In order to have simulation results close to what one would observe in real world interactions, data employed in the simulation is from the data collection of the study in [168]. In addition, while the display of emotional expression was not studied, based on literature review in Chapter 2, this study assumes that if the robot displays appropriate cooperative *emotional expressions*, it would increase the likelihood of the human to take its suggestion and thus have a $D = +1$ in the reward function (Equation 6.2).

The simulated human was designed as follows and shown in Figure 6.3 with a distribution of 100 samples and all claims made are based on findings in [168]:

1. If the robot exhibits proxemics, the user would significantly stand closer to it, as shown in Figure 6.3a.
 - In the *presence* of proxemics: the distance of the user is selected randomly from a normal distribution that has $\mu_{dh} = 0.801$ meters and $\sigma_{dh} = 0.112$ and the minimum distance dh was 0.51 meters.
 - In the *absence* of proxemics: the distance of the user is selected randomly from a normal distribution that has $\mu_{dh} = 1.03$ meters and $\sigma_{dh} = 0.216$ and the minimum distance dh was 0.71 meters.
2. If the robot displays proxemics behavior and/or emotional expressions, the user would be more likely to take the recommendation.
 - In the *presence* of robot's proxemics behavior and/or emotional expressions, the user has a 65% probability of taking the robot's recommendation and 35% probability of not following the robot's suggestion.

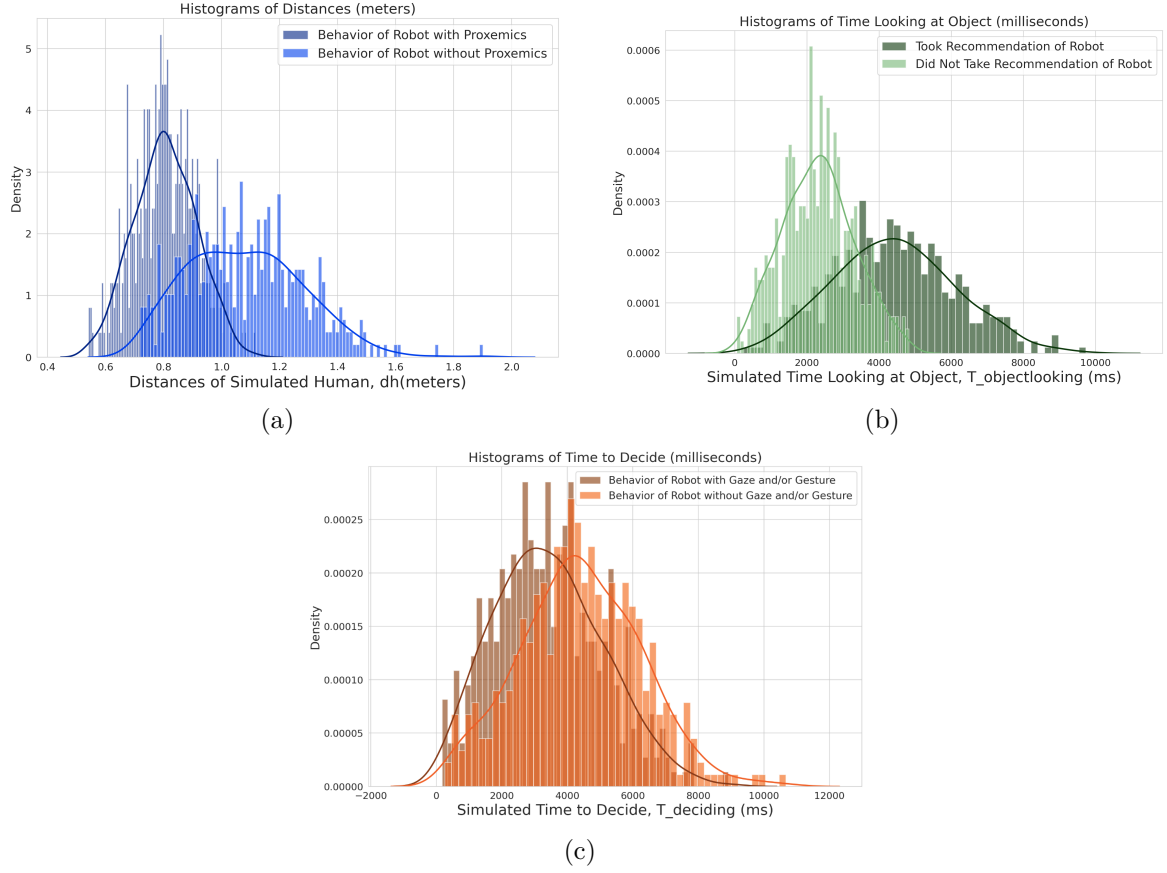


Fig. 6.3 Histograms of data distributions for the simulation. Figure(a) shows the normal distributions of the distances held by the user dh (meters) when the robot's behavior includes proxemics and when the robot's behavior does not include proxemics. Figure(b) shows the normal distributions of the time spent looking at the referred object when the user takes the recommendation of the robot and when the user does not. Figure(c) shows the normal distributions of the time the human spent making a decision when the robot's behavior includes gaze and/or gesture and when the robot's behavior does not include gaze nor gesture.

- In the *absence* of robot's proxemics behavior and/or emotional expressions, the user has a 20% probability of taking the robot's recommendation and 80% probability of not following the robot's suggestion.
3. If the user took the recommendation of the robot, they would more likely spend a longer amount of time looking at the referred object by the robot ($T_{objectlooking}$), as shown in Figure 6.3b.
 - If the simulated human followed the suggestion of the robot, $T_{objectlooking}$ is randomly selected from a normal distribution of $\mu_{T_{objectlooking}} = 4452.77$ ms and $\sigma_{T_{objectlooking}} = 1714$ and having only positive values.

- If the simulated human *did not* follow the suggestion of the robot, $T_{objectlooking}$ is randomly selected from a normal distribution of $\mu_{T_{objectlooking}} = 2164.24$ ms and $\sigma_{T_{objectlooking}} = 1032.25$ and having only positive values.
4. If the robot displays gestures and/or gaze behaviors, the user would more likely spend a shorter amount of time deciding ($T_{deciding}$), as shown in Figure 6.3c.
- In the *presence* of robot's gaze and/or gesture behaviors, $T_{deciding}$ of the user is randomly selected from a normal distribution of $\mu_{T_{deciding}} = 3111$ ms and $\sigma_{T_{deciding}} = 1817$ ms and having only positive values.
 - In the *absence* of robot's gaze and/or gesture behaviors, $T_{deciding}$ of the user is randomly selected from a normal distribution of $\mu_{T_{deciding}} = 4314.6$ ms and $\sigma_{T_{deciding}} = 2026$ ms and having only positive values.

6.4 Implementation of Reinforcement Learning Models

k	$k \in [0, 3]$	$k \in [4, 9]$	$k \in [10, 13]$	$k \in [14, 15[$
	one-modality	two-modalities	three-modalities	four-modalities
k: Multi-Modal Behavior	0: <i>Proxemics</i>	4: <i>Proxemics + Gaze</i>	10: <i>Proxemics+Gaze + Gesture</i>	14: <i>Proxemics + Gaze + Gesture + Emotional Expressions</i>
	1: <i>Gesture</i>	5: <i>Proxemics+ Gesture</i>	11: <i>Proxemics+Gaze+ Emotional Expression</i>	
	2: <i>Gaze</i>	6: <i>Proxemics +Emotional Expression</i>	12: <i>Proxemics+ Gesture+Emotional Expression</i>	
	3: <i>Emotional Expression</i>	7: <i>Gesture+ Gaze</i>	13: <i>Gesture+Gaze+ Emotional Expression</i>	
		8: <i>Gesture+Emotional Expression</i>		
		9: <i>Gaze+Emotional Expression</i>		

Table 6.1 $k = 15$ representing $[0, k[$ behaviors of varying degrees of modality

6.4.1 Multi-Armed bandit

Multi-Armed Bandits, or *k-Armed Bandits*, were implemented as a first step to investigate the weights of the reward function and its different parts as well as to serve as a baseline algorithm. Moreover, it was a simple method to have a first look into the multi-modal combinations different agents with varying ϵ 's and learning rates α 's might select. To do so, $k = 15$ was chosen throughout the simulation to represent the number of possible combinations that can exist for the chosen modalities ($k = 15$ implies $k \in [0, 15[$. k from $[0, 3]$ represent one-modality behaviors, $[4, 9]$ represent two-modality combined behaviors, $[10, 13]$ represent three-modality combined behaviors, and 14 represents the four-modality combined behavior. The k-actions representations are shown in Table 6.1.

6.4.2 Q-Learning

The *Q-Learning* algorithm and model was implemented to better suit the problem of adaptation and personalization in a HRI set-up, as shown in Algorithm 2. First, an ϵ -Decay strategy for the ϵ -greedy action selection was used to fill the Q-table rather than a fixed value ϵ . This was done so to allow high exploration in the beginning before converging into exploitation. At first, the agent does not have information about the environment, in this case the human, it is interacting with. Only once the agent has the information it needs to interact optimally with the human/environment, it can better exploit its knowledge and adapt more appropriately. The ϵ decayed after every episode as shown in Equation 6.3:

$$\epsilon = \epsilon * e_{decayrate} \quad (6.3)$$

where $e_{decayrate} = 0.9998$ and the start value of $\epsilon = 0.9$. After running through a large number of episodes, filling the Q-table, and letting ϵ decay to a smaller number, the filled Q-table can then be used to rerun the algorithm with a fixed smaller ϵ this time to optimally interact with the human while keeping a small window for exploration.

Moreover, the state-space and action-space were implemented as described in Subsection 6.2.

Algorithm 2: Q-Learning with Adaptive Reward Function and ϵ -Decay

```

Initialize action-value function  $Q(s, a)$  with random values  $\in [0, 1[$  ;
foreach episode  $e \leftarrow 1$  to  $M$  do
    Initialize state  $s_t$ ;
    foreach step  $t \leftarrow 1$  to  $T$  do
        With probability  $\epsilon$  select a random action  $a_t$  otherwise select  $a_t = \max_a Q(s_t)$  ;
        Execute action  $a_t$  to observe state  $s_{t+1}$ ;
        Get current reward from observed state  $r(s_{t+1})_t$  ;
         $r(s_{t+1}) \leftarrow -\kappa(d_h) + \tau \left( \frac{T_{objectlooking}}{T_{deciding}} \right)^{\frac{1}{\phi}} + \beta(D) - \delta(degree_{modality})$  ;
        Update Q(s,a) values:
             $Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha(r(s_{t+1}) + \gamma \max_a Q(s_{t+1}, a_t))$ ;
     $\epsilon \leftarrow \epsilon * e_{decayrate}$ ;

```

6.5 Results

6.5.1 Multi-Armed Bandit

In this section, several multi-armed bandit algorithms with ϵ -greedy action selection (ϵ -MAB) and upper-confidence bound with ϵ -greedy action selection (ϵ -UCB) algorithms with varying ϵ , c , and learning rate α were used. This was done so to better understand and evaluate the outcome for different exploitation-exploration balance, upper-confidence bound exploration, and learning rate respectively. Four ϵ -MAB algorithms were used: $\epsilon = 0.1$ with $\alpha = 0.1$, $\epsilon = 0.1$ with $\alpha = 0.5$, $\epsilon = 0.5$ with $\alpha = 0.1$, and $\epsilon = 0.5$ with $\alpha = 0.5$. In addition, four ϵ -UCB algorithms with a fixed learning rate of $\alpha = 0.1$ were implemented: $c = 2$ with $\epsilon = 0.1$, $c = 2$ with $\epsilon = 0.5$, $c = 5$ with $\epsilon = 0.1$, and $c = 5$ with $\epsilon = 0.5$.

First, the reward function with only the social signals part was used to evaluate the effect it would have on the most and least selected multi-modal behaviors by the algorithms. Second, the full reward function, shown in Equation 6.2, with two different δ values were implemented and the algorithms were ran to study the effect of varying weights in the cost function has on the multi-modal behaviors selected. The outcomes investigated for each algorithm were their learning curves through the moving average episode reward and, in order to highlight the adaptation through the combinations of modalities chosen by the robot, the percentage of times each action, or k as shown in Table 6.1, was selected at various iteration rates for 20 episodes with different random seeds. Additionally, these evaluations help better understand the impact of the weights in the reward function and would later serve to design and implement a baseline algorithm.

Social Signal Reward

Before investigating the full reward function, a reward function using only the human’s social signal parts was evaluated using the four ϵ -MAB and four ϵ -UCB algorithms. The social signals reward function used is shown in Equation:

$$R_{socialsignal} = -\kappa(d_h) + \tau \left(\frac{T_{objectlooking}}{T_{deciding}} \right)^{\frac{1}{\phi}} \quad (6.4)$$

where $\kappa = 1.587$, $\tau = 1.27$, and $\phi = 2$. The weights were fine-tuned to have a balanced impact from both parts of the $R_{socialsignal}$ function.

All algorithms were ran using the social signals reward function for 100 and 5000 iterations. The average episode social signal rewards for 100 and 5000 iterations are shown in Figure 6.4. At 5000 iterations, the algorithms have already converged with ϵ -MAB ($\epsilon = 0.1, \alpha = 0.1$) yielding the highest reward. In addition, the multi-modal behaviors or k ’s selected by each algorithm were averaged over the 20 episodes with different random seeds and their percentages are represented in Figure 6.5 after 100 iterations and figure 6.6 after 5000 iterations.

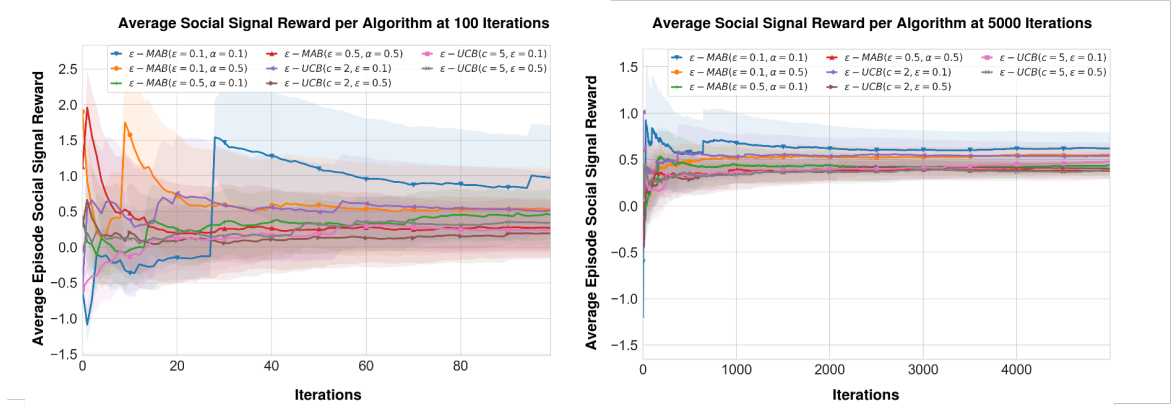


Fig. 6.4 The learning curves showing the episodic average social signal reward per algorithm after 100 (left) and 5000 (right) iterations per 20 random seeds episodes respectively. The mean and standard deviation over the 20 random seeds are plotted. The reward function used contained only the human’s social signals part.

At 100 iterations, ϵ -MAB algorithms already favor high degree modalities, i.e. three-modality, except the three-modality behavior *gaze, gesture, and emotional expressions*, four-modality behaviors, as well as the two-degree modality behaviors, proxemics with gaze behaviors ($k = 4$) and proxemics with gesture behaviors ($k = 5$). On the other hand, at 100 iterations, the ϵ -UCB select approximately all multi-modal behaviors (k -actions) at the same percentage. After 5000 iterations and convergence of the algorithms, all algorithms

favor selecting the three-modality behaviors ($k \in [10, 11]$), except the three-modality behavior *gaze, gesture, and emotional expressions*, the four-modality behavior ($k = 14$), and the two two-modality behaviors (*proxemics with gaze behaviors* ($k = 4$) and *proxemics with gesture behaviors* ($k = 5$)). With algorithm ϵ -MAB ($\epsilon = 0.1, \alpha = 0.1$) yielding the highest reward, most likely due to its exploitation rate, the multi-modal behavior most selected was two-modality (*proxemics with gaze*) behavior, followed by three-modality behavior (*proxemics, gaze, and emotional expression*), and tying in third place are two-modality behavior (proxemics with gesture) and three-modality behavior (proxemics, gesture, and emotional expressions). The least selected behaviors by ϵ -(MAB) ($\epsilon = 0.1, \alpha = 0.1$), ϵ -(MAB) ($\epsilon = 0.1, \alpha = 0.5$), and ϵ -(UCB) ($c = 2, \epsilon = 0.1$) are one-modality behavior (Emotional Expressions), two-modality behaviors (*Gesture with Emotional Expression* and *Gaze with Emotional Expression*), and three-modality behavior *gaze, gesture, and emotional expressions*.

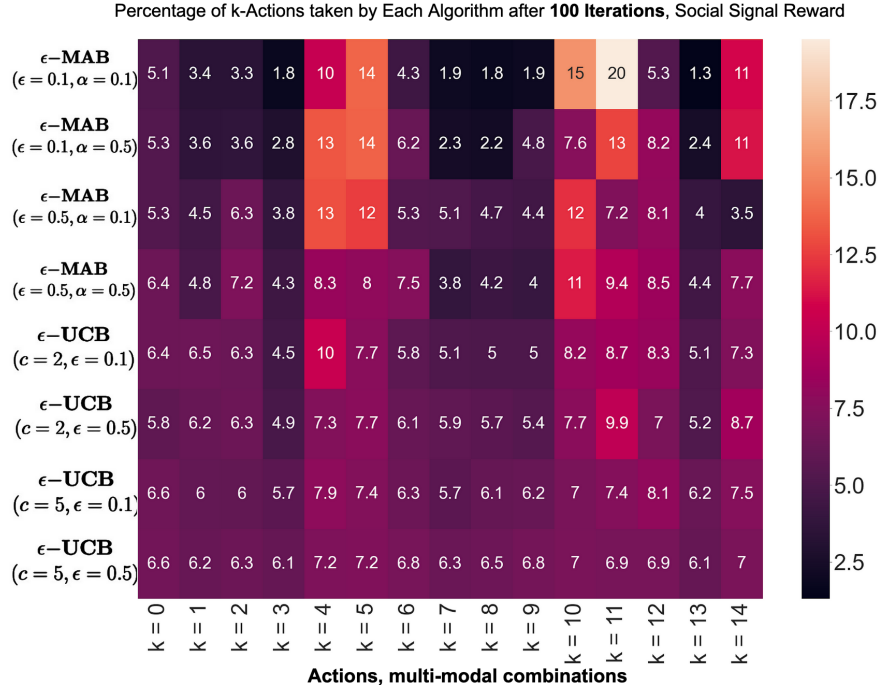


Fig. 6.5 Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 100 per 20 episodes with the social signal reward respectively

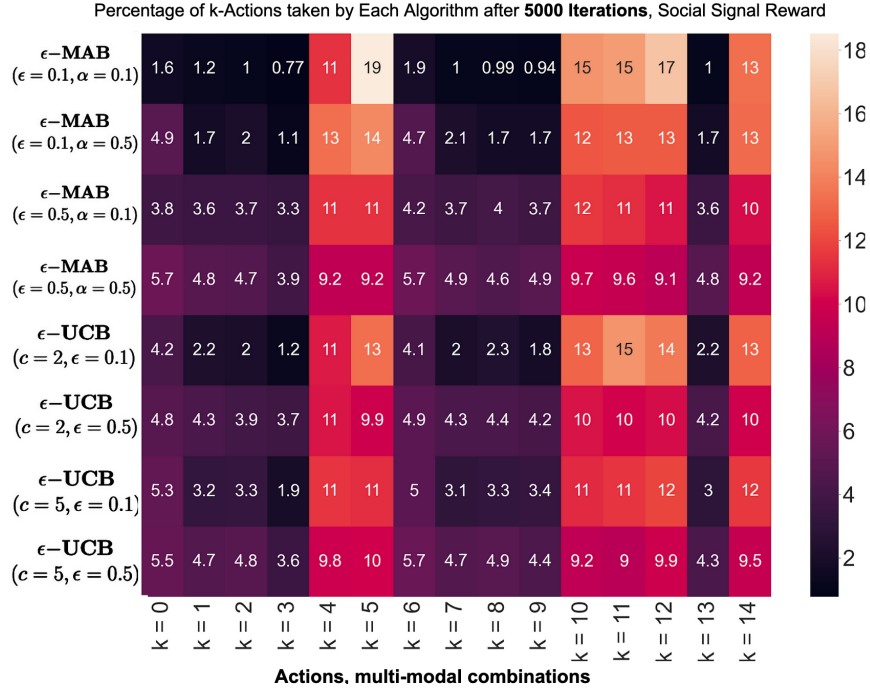


Fig. 6.6 Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 5000 iterations per 20 episodes with the social signal reward respectively

Full Reward

In this subsection, the entire reward function of Equation 6.2, which takes into account the human social signals and their decisions as well as the robot's cost of high-degree modality behaviors, was used and all algorithms were run for 100, 1000, and 10000 iterations. The weight value of β was set to 0.501. Two different values for the delta weight were examined to further investigate the impact of the weights of the cost function on the multi-modal behavior choices of the robot.

First, $\delta = 0.24$ was set and the learning curves of all algorithms are shown in Figure 6.7. At 10000 iterations, the convergence can be seen and the algorithm ϵ -MAB ($\epsilon = 0.1, \alpha = 0.1$) yielded the highest reward. In addition, the percentage of selection for each k, which represent the multi-modal behaviors are plotted in Figure 6.8 at several instances: after 100 iterations (Figure 6.8a), 1000 iterations (Figure 6.8b), and 10000 iterations (Figure 6.9). At 100 iterations, the algorithms had not yet converged and most of which have an equally distributed percentage selection for the k-actions (multi-modal behaviors). However, even at this early point, for ϵ -MAB ($\epsilon = 0.1, \alpha = 0.1$) the most selected multi-modal behaviors are the two two-modality behaviors: *proxemics with gaze* and *proxemics with gesture*. Moreover, at already 1000 iterations, the previously mentioned two-modality behaviors become among

the most selected along with one-modality behavior made of only *proxemics*. Finally at 10000 iterations and convergence, these three multi-modal behaviors (*proxemics with gaze*, *proxemics with gesture*, and *proxemics* alone) become the most favored for all algorithms. As an effect of the cost function, higher degree modality behaviors are least selected, especially the four-modality behavior.

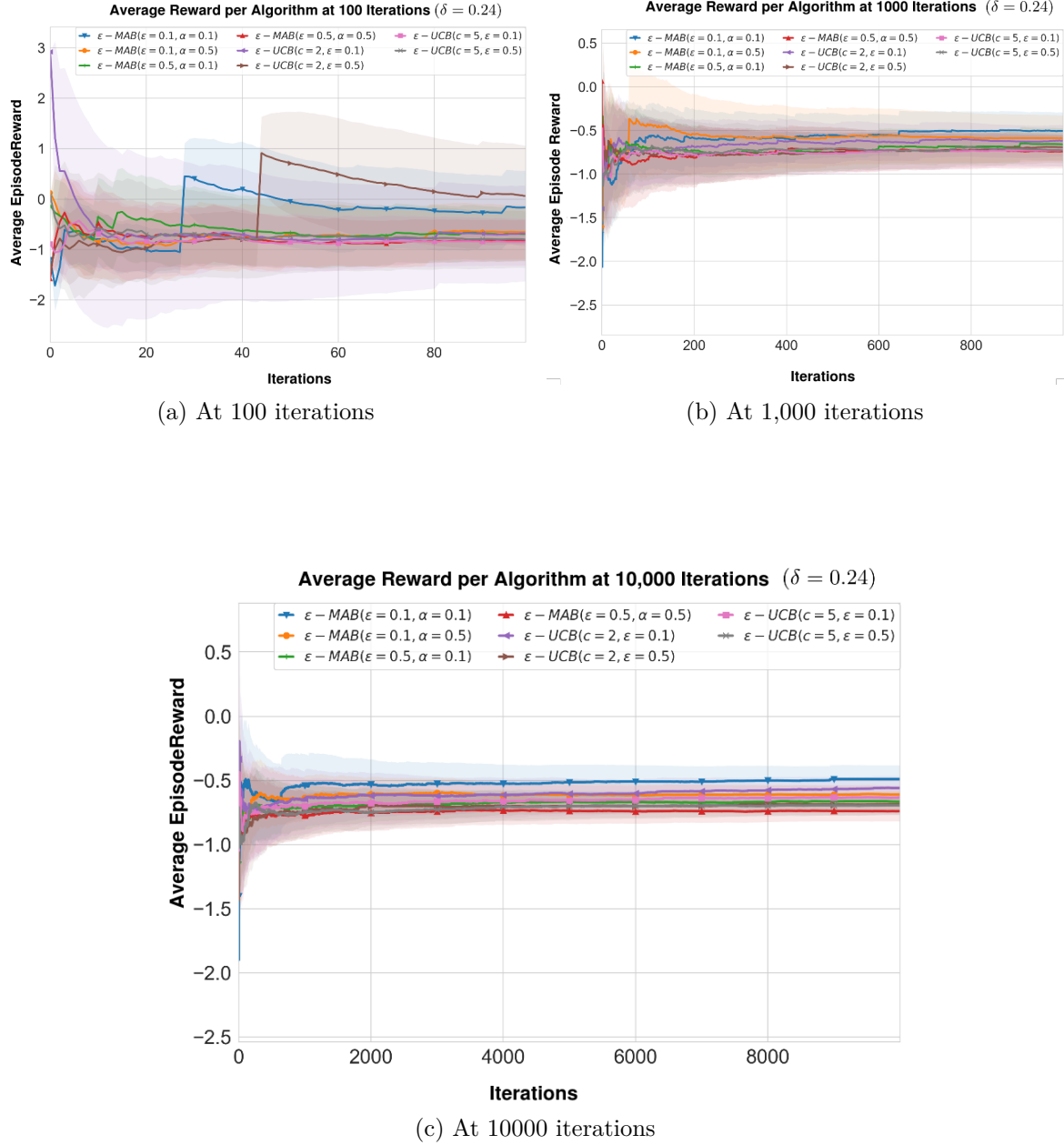


Fig. 6.7 Average episodic reward per algorithm after (a) 100 (b) 1,000 and (c) 10000 iterations for 20 random seeds episodes with the full reward respectively($\delta = 0.24$). The mean and standard deviations are plotted.

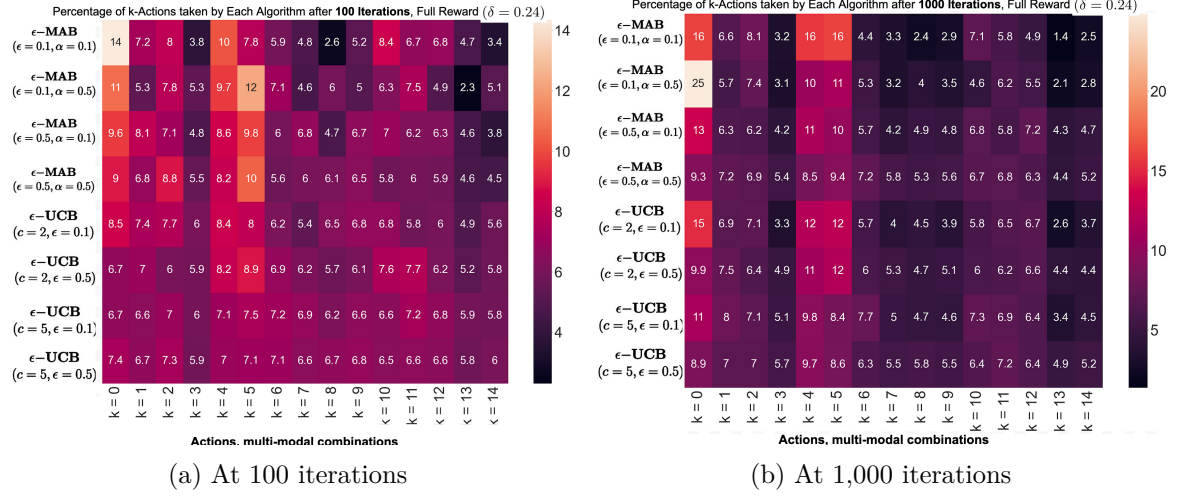


Fig. 6.8 Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after (a) 100 and (b) 10000 iterations per 20 episodes with the full reward respectively ($\delta = 0.24$). The average percentage of the 20 random seeds is shown.

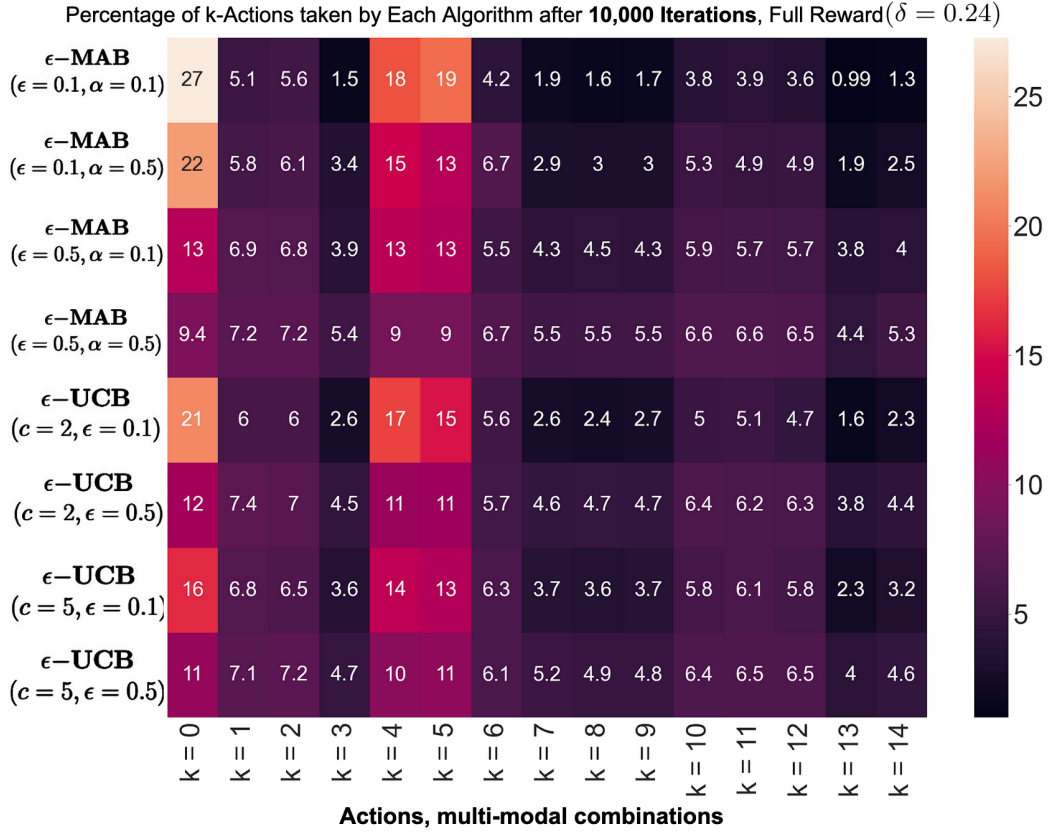


Fig. 6.9 The percentage of k-actions (representing multi-modal behaviors) taken by each algorithm after 10000 iterations per 20 episodes with the full reward respectively ($\delta = 0.24$). The average percentage of the 20 random seeds is shown.

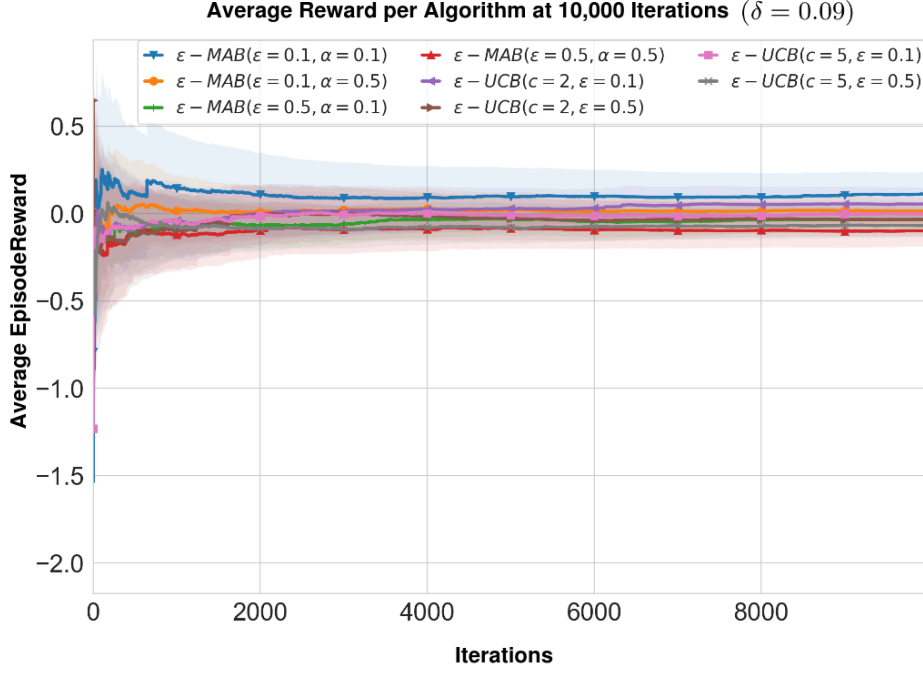


Fig. 6.10 Average reward per episode of algorithms using full reward with weight $\delta = 0.09$ after 10000 iterations for 20 random seed episodes

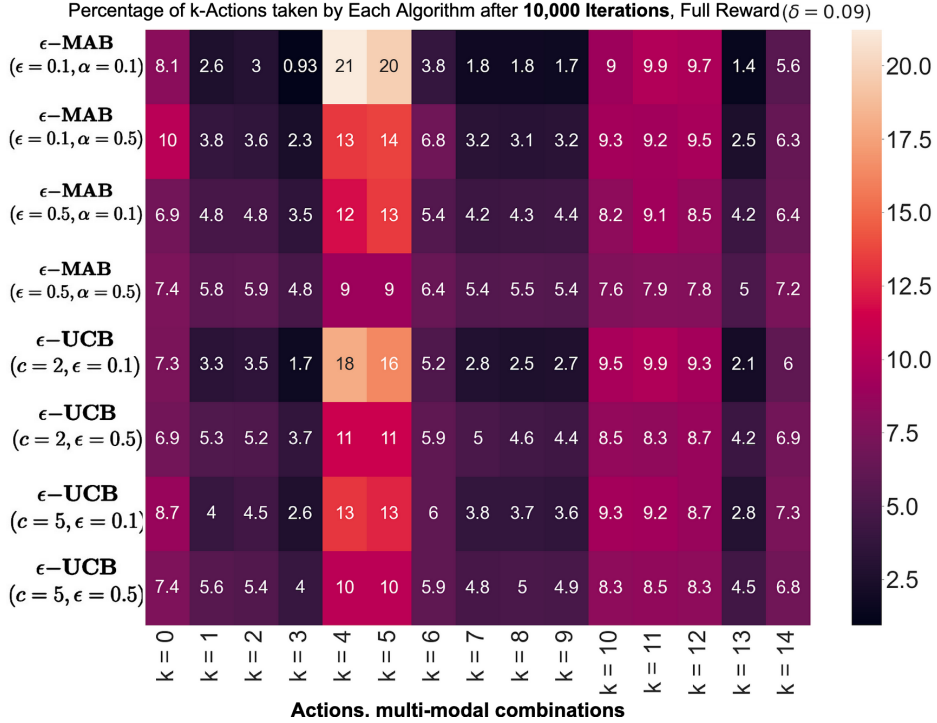


Fig. 6.11 Percentage of k-actions (representing multi-modal behaviors) taken by each algorithm using full reward with weight $\delta = 0.09$ after 10000 iterations for 20 random seed episodes

Second, in order to not let the cost function limit exploration for the agent while maintaining the trade off between optimizing behaviors for human's social signals and the cost high-degree modality behaviors may have on the robot, the weight δ was reduced to 0.09. All algorithms were run for 10000 iterations and the learning curve is shown in Figure 6.10. The percentage of selected k-actions (multi-modal behaviors) by each algorithm are shown in Figure 6.11. Decreasing the value to $\delta = 0.09$ allowed for three-modality behaviors to be more selected than when $\delta = 0.24$ was set, except for the three-modality behavior *gaze, gesture, and emotional expressions*, as well as slightly increasing the percentage of selection for the four-modality behavior. The most selected behaviors remain the two-modality behaviors: *proxemics with gaze* and *proxemics with gesture*. However, in comparison to Figure 6.9 with $\delta = 0.24$, the selection of other multi-modal behaviors with $\delta = 0.09$ are more evenly distributed giving more room for exploration. The least selected behaviors for $\delta = 0.09$ are the the two-modality behaviors: *gesture with gaze*, *gesture with emotional expressions*, *gaze with emotional expressions*, and three-modality behavior *gaze, gesture, and emotional expressions*, in addition to the modality *emotional expressions* on its own.

6.5.2 Q-Learning

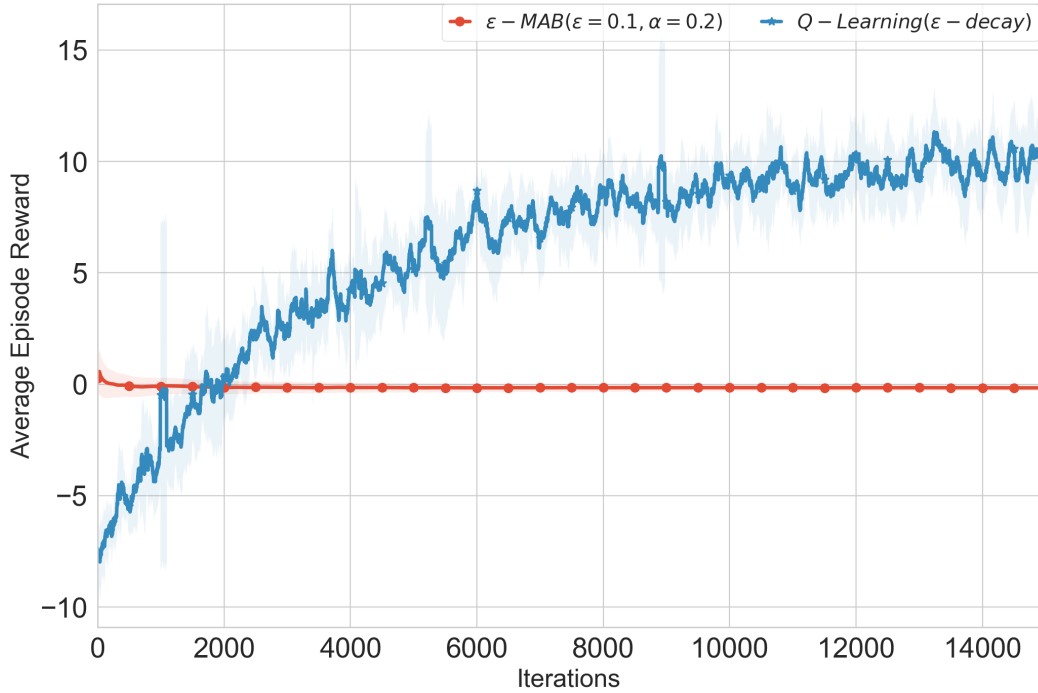


Fig. 6.12 Average episode reward after 15000 episodes for 10 random seeds. The means and standard deviations are plotted. ϵ -MAB($\epsilon = 0.1, \alpha = 0.2$) serves as a baseline.

The Q-Learning algorithm was implemented as shown in Algorithm 2 resulting in a Q-Learning (ϵ -decay). The weights for the reward function of Equation 6.2 were set similarly to the ϵ -MAB and ϵ -UCB algorithms. In addition, δ was set to 0.09 to not restrict exploration of multi-modal behaviors caused by the cost function. The learning rate was fixed at $\alpha = 0.2$ and the discount factor at $\gamma = 0.95$. The initial epsilon was set at $\epsilon = 0.99$ and the decay rate at $e_{decay} = 0.9998$. In addition, each episode would run 50 steps while updating the Q-table. The number of steps were chosen to decrease the noise in the reward signal and increase the average episode rewards.

The algorithm was ran for 15000 episodes for 10 random seeds and the learning curve of the mean with standard deviation is presented in Figure 6.12. The value of ϵ throughout decreased from 0.99 to 0.05 as the average episode reward curve converges. ϵ -MAB ($\epsilon = 0.1, \alpha = 0.2$) with the same reward function was ran and serves as a baseline. Q-Learning (ϵ -decay) yielded a much higher average episode reward than the baseline though it took longer to converge.

Furthermore, the Q-table of $Q(s, a)$ values is mapped up in Figure 6.13. For all states where proxemics is absent, the $q(s, a)$ values are highest for the action *proxemics* = +1 while all other actions for the same states have the lowest values of $q(s, a)$. On the other hand, for all states where proxemics is present, the $q(s, a)$ are more distributed between the actions. For state (1, 0, 0, 0) (*proxemics* only), the action with the highest $q(s, a)$ value is for action *gaze* = +1.

In addition, the percentage each state was visited is shown in Figure 6.14. The states most visited were the two-modality behaviors: *proxemics with gaze* and *proxemics with gesture*, followed by three-modality behavior: *proxemics with gaze and gesture*, and then three-modality behaviors: *proxemics with gaze and emotional expressions* and *proxemics with gesture and emotional expressions* as well as one-modality behavior with *proxemics* alone. On the other hand, the state least visited was the behavior with no multi-modality followed by all states that do not include proxemics.

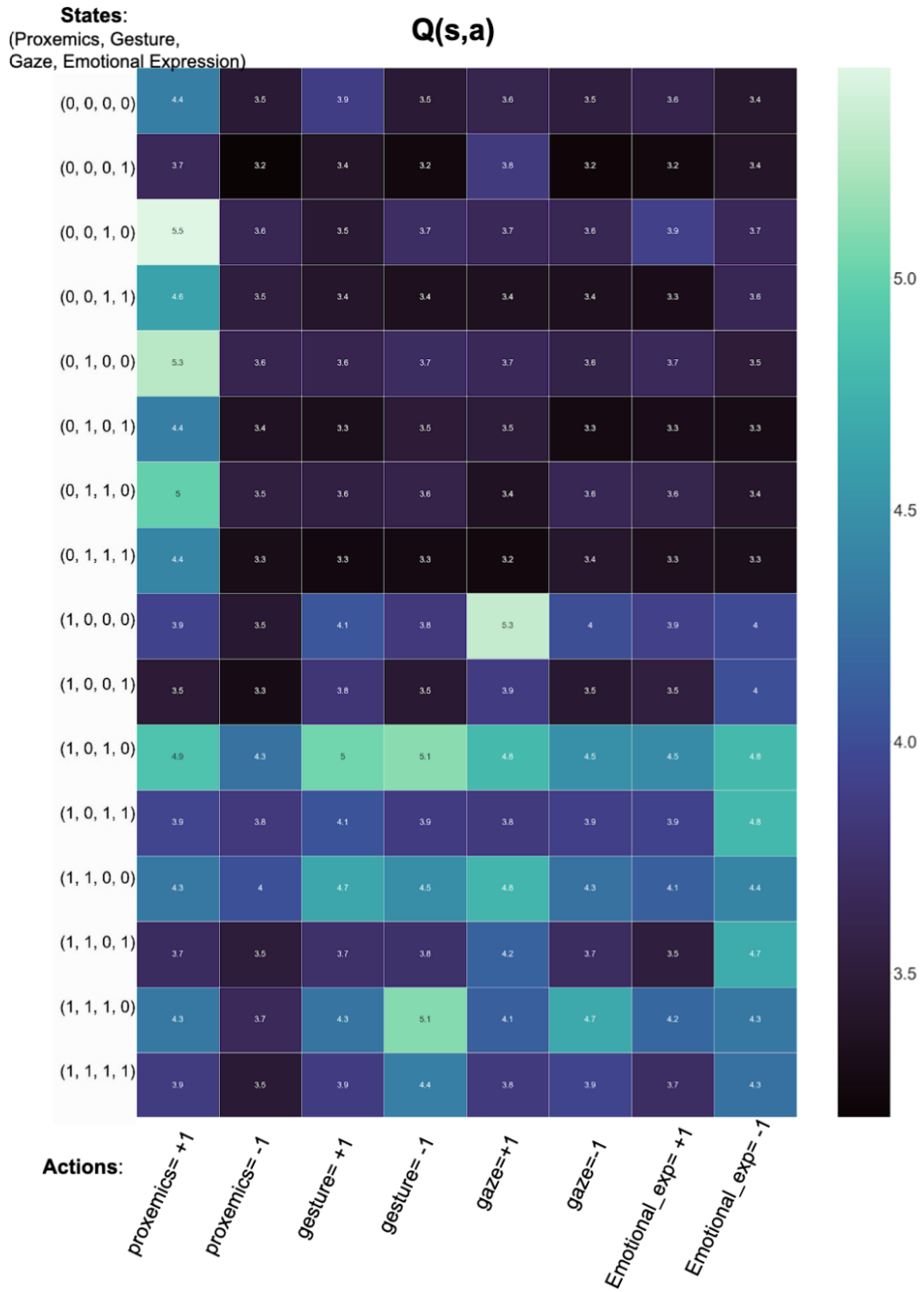


Fig. 6.13 Q-Table with the average $q(s, a)$ values after 15000 episodes for 10 random seeds for Q-Learning(ϵ -decay) algorithm..

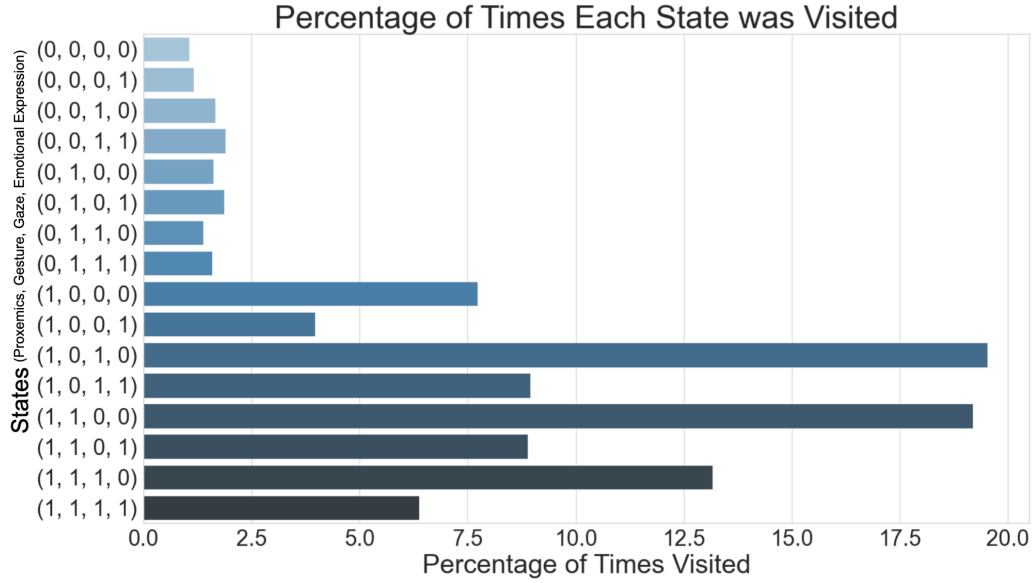


Fig. 6.14 Average percentage of states visited after 15000 episodes for 10 random seeds for Q-Learning (ϵ -decay) algorithm.

Further evaluation was done to see the impact the RL models had on the simulated human behaviors and the human's social signals. The results are plotted in Figure 6.15. First, Figure 6.15a shows the average simulated human's distance in meters. With the Q-Learning(ϵ -decay) model, the simulated human stood closer to the robot after about 4000 episodes as the robot was learning. Second, Figure 6.15b shows the average of times the simulated human took the suggestion of the robot, where a value of 1 signifies that the human took the robot's suggestion and 0 the human did not. With the Q-Learning(ϵ -decay) model, the human took the robot's suggestion more throughout all 15000 episodes. Third, Figure 6.15c shows the average time the simulated human spent looking at the object referred to by the robot in milliseconds. With the Q-Learning(ϵ -decay) model, the time spent looking was higher throughout all 15000 episodes. Fourth, Figure 6.15d shows the time spent by the simulated human deciding whether to take robot's suggestion or not (in milliseconds). With the Q-Learning(ϵ -decay) model, time spent deciding was lower than the baseline after about 4000 episodes and until the end of the 15000 episodes.

In addition, the time it takes for both Q-Learning(ϵ -decay) algorithm and baseline to run the learning episodes needs to be considered. The average time it took to run one episode in Q-Learning(ϵ -decay) was 211.24 seconds with a standard deviation of 7.07 seconds. On the other hand, the average time it took to run one episode in the baseline ϵ -MAB algorithm was 65.61 seconds with a standard deviation of 7.25 seconds.

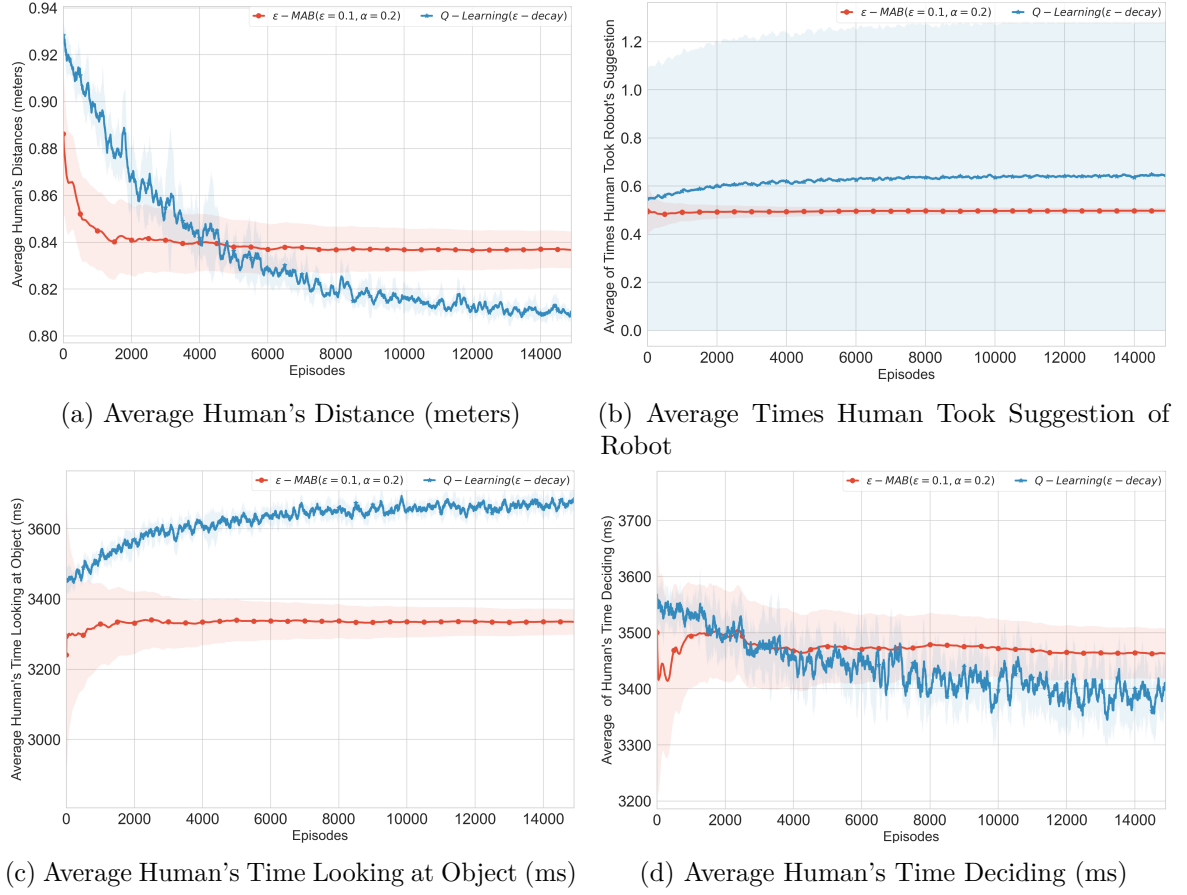


Fig. 6.15 Average Simulated Human's Distance(meters), Average Times Human Took Suggestion of Robot, Average Simulated Human's Time Looking at Object (ms), and Average Simulated Human's Time Deciding (ms), after 15000 episodes of 10 random seeds, by each algorithm: ϵ -MAB($\epsilon = 0.1, \alpha = 0.2$) and Q-Learning(ϵ -decay). The mean and standard deviations are plotted.

6.6 Discussion

In HRI, human's social signals provide important information and the robot's functional multi-modal behavior allow it to effectively communicate and be perceived as a socially intelligent agent. The continuous nature of human's social signals was used in the reward function to avoid discretizing them and losing information as well as to decrease the size of the state-space and consequently increase the learning time. In addition, this chapter presented simulation of HRI for training of algorithms and this is an important contribution to the field of HRI specially during the COVID-19 pandemic, which made physical data collection very difficult and as such slowing down many studies. Furthermore, the findings in the result section may provide important lessons and guidelines for adaptation and personalization in

HRI using reinforcement learning.

First, the results from running the ϵ -MAB and ϵ -UCB algorithms with only the social signals part of the reward function showed that the high-degree modalities were favored, except for the three-modality behavior *gaze*, *gesture*, and *emotional expressions*. High modality behaviors contain all modalities needed for rich social interactions. However, the algorithms selected just as much the two two-modality behaviors, *proxemics with gaze* and *proxemics with gesture*. This demonstrated that these two-modality behaviors may be as effective and rich in social interactions as the higher-degree modalities. In addition, the only three-degree modality not highly selected was the one without proxemics. Already at this point, proxemics may play an important role for the behavior of robots.

Second, the previous finding was also highlighted when using the full reward function with the weight of the cost function on the degree of modality set as high as $\delta = 0.024$. The cost function caused the algorithms to disregard high-degree modality behaviors (three-modality and more) and select much more the two two-modality behaviors (*proxemics with gaze* and *proxemics with gesture*) as well as *proxemics* alone, which also had highest percentage of selection by some of the algorithms. This emphasized the value and influence of the proxemics behavior in social robots when compared to other modalities, specially as a modality that can interact efficiently alone. Furthermore, proxemics is a modality that is inherent to robots in comparison to other socially intelligent machines such virtual agents. Robots are able to navigate and re-orient themselves by spotlighting that their embodiment allows them to be in the same physical environment as the human.

Third, when δ was decreased to 0.09, the restrictions on high degree modality behaviors was relaxed and increased the percentage of selecting them. The two most selected behaviors remain the two two modality behaviors (*proxemics with gaze* and *proxemics with gesture*). However, this time the percentage for selecting the one-modality behavior of *proxemics* alone and the three three-modality behaviors (*proxemics, gaze, and gesture*, *proxemics, gaze, and emotional expression*, and *proxemics, gesture, and emotional expressions*) were approximately the same. This would allow the robot to explore more behaviors while still having a cost on the degree of modalities making up its behavior.

Fourth, keeping $\delta = 0.09$, the Q-Learning(ϵ -decay) and ϵ -MAB, which was a baseline, were ran for 15000 episodes. The average episode reward by Q-Learning(ϵ -decay) was lower than the baseline before reaching 2,000 episodes as it was a period of very high exploration. After 2,000 episodes, as the number of episodes increased and the ϵ decayed, Q-Learning(ϵ -decay)'s

average episode reward reached a significantly higher value than the baseline. Similarly to the previously ran algorithms, after 15000 episodes, the most visited states, which represent the multi-modal behaviors of the robot, are those that include proxemics. On the other hand, the least visited state (behavior) was the one with *no modalities*. This also highlights the value of modalities in improving the quality of interactions.

The two most visited states are the two two-modality behaviors, *proxemics with gaze* and *proxemics with gesture*. Gesture and gaze play several functional roles such as referring to objects in the environment of interaction and communicating internal processes such as turn-taking, floor-holding, and turn-yielding in conversations as well as expressing a listening state through back-channeling. In [168], it was shown that gestures significantly increased user's behavioral alignment with the robot by waving at it at the start and end of the interaction. In addition, gaze significantly increased the likelihood of users performing backchanneling while the robot was talking.

Fifth, the heat-map of the Q-table also reflected the importance of proxemics since the action with the highest value for states with no proxemics was to activate proxemics. As for the rest of the states, the values of all of the actions were approximately equally distributed. This allows for further combinations of modalities to be formed and for exploration, which is also needed for adaptation and personalization along with balanced exploitation. The Q-table provided by Q-Learning(ϵ -decay) after training on the simulated human and interaction is important as it can be used as an initial Q-table when the robot is to interact in real life with human. Having an already trained and filled Q-table would allow for faster learning and as such a faster adaptation.

Finally, when looking into the simulated behaviors of the human, the findings showed that the human behaved in a more desired manner with the Q-Learning(ϵ -decay) algorithm than with the baseline. After 15000 iterations, the simulated human was standing closer to the robot, specially at a more social distance of approximately 0.81 meters, taking the robot's suggestions more, spending more time looking at the object referred to by the robot, and taking less time deciding to take the robot's suggestion. Moreover, the weights in the reward function have shown that by increasing or decreasing their values, one can slightly drive the exploration in one direction or another and as such further exploration of combinations of modalities making up the robot's behavior can be investigated.

6.7 Limitations

While this study has put forward interesting findings, it has certain limitations that need to be addressed. First, the proposed RL model in this chapter has not been tested yet in real-world environments. Future work includes implementing the model on a physical robot and collecting data with humans to better investigate the adaptation, multi-modal behaviors selected by the robot, and the learning of the robot. Second, the social signals of humans are very complex and coupled in nature and while the reward function presented captures some, it does not capture all. There may be more terms that can be added to the human social signals part of the reward function. Third, emotional expressions may play a bigger role for instance on the decision-making of humans and their responsive social signals. Moreover, emotional expressions are themselves multi-modal behaviors in nature. This study has only relied on findings from literature review, but more investigation and observation needs to be done to better understand its impact and have its weight in the reward function adjusted accordingly.

6.8 Conclusion

This chapter presented a simulation set-up of human-robot interactions and a RL model for adaptation and personalization of interactions taking into account the human social signals, human decisions, and the cost of the complexity of behaviors. The findings have allowed us to learn the most influential modalities and combinations of multi-modal behaviors. For instance, the findings highlighted the imperative need of proxemics for successful interactions with embodied robots. Moreover, the equivalence between gaze and gesture for their functionalities was brought into light. Such as, different situations may require and prefer one over the other. To conclude, this chapter demonstrated a general adaptation, which can be used in real-life interactions by making use of the resulting Q-table, specially since it has shown potential for personalization for different user profiles and situations and it has illustrated the value of RL in adaptation.

Chapter 7

Discussion & Conclusion

This chapter presents a discussion and conclusion to the work described in this doctoral dissertation. It summarizes the key research findings in relation to the research aims and questions and discusses the value and contribution thereof. It will also review the limitations of the study and propose opportunities for future research.

7.1 Discussion

7.1.1 Thesis Overview & Contributions

The goal of this thesis is to investigate and synthesize multi-modal socially intelligent human-robot interaction. For an interaction to be perceived as socially intelligent, the robot needs to be able to assess and manage individuals it is interacting with and respond through its capability of using and adapting its behavior. In addition, robots are embodied agents that can communicate non-verbally and manipulate the environment around them using their multi-modal behaviors. This thesis aimed at contributing to the field of HRI by designing and implementing autonomous multi-modal behavioral solutions for a robot to investigate their impact on the users interacting with it and presenting a RL model that allows the robot to appropriately adapt and personalize its behavior by taking into account human social signals. To demonstrate the findings, the thesis was divided into two parts: *Relative Work* and *Contributions*.

The *Relative Work* part focused on discussing the state-of-the-art and literature review. Chapter 2 introduced the studies done investigating the modality behaviors of robots including proxemics, gaze, gesture, and emotional expressions and the link between behavior and social intelligence and influence. While there has been many contributions on designing and

understanding modalities individually, their impact when performed in combination in a multi-modal behavior is still not clear. Moreover, there is a gap in apprehending the extent of how human multi-modal behaviors can be used by the robot to manipulate social situations. Thus, there is a need to develop a comprehensive perspective on how different combinations of these modalities can impact and influence interactions. Chapter 3 presented a background on RL and an in-depth literature review on RL in social robotics for adaptation. The recent work have shown the potential of RL to solve for adaptation, but so far the exploration of adaptation by combining modalities to form different multi-modal behaviors while taking into account the continuous social signals of humans and the cost of complexity put on the robot has not been achieved.

The *Contributions* part of the thesis focuses on addressing and contributing to the main research question through three main studies: *"When found in an interaction set-up, how should the robot use its multi-modal behavior and understanding of the interaction dynamics with the humans around it to act as a socially intelligent agent?"*. The three main studies were presented in Chapter 4, which studied how one human modality may be used by the robot to adapt its behavior to different situations, Chapter 5, which synthesized multi-modal behaviors of the robot and investigated their effects on the outcomes of the interaction, and Chapter 6, which presented a simulation setup of HRI for RL training and a RL model to allow the robot to adapt its multi-modal behavior to human social signals through the reward function.

RL for adaptation in HRI: An In-Depth Literature Review

Using RL as an approach to adaptation in HRI is relatively new and there is still a need to further investigate it and answer the gaps and challenges present. However, in order to do so, it is crucial to first dive into the literature of the recent RL models of adaptation in social robotics. With a lack of such a literature review, the first contribution of the thesis is to present an in-depth study on the RL approaches used for adaptation in HRI by demonstrating the details of the proposed RL models, the findings of the studies, and their limitations. One of the limitations is found in the studies that attempted to discretize human social signals and having it lead to large state spaces [17, 139, 183]. Moreover, most of the studies focused only on a one-modality behavior [117, 96] or only verbal responses [17, 58, 105]. In addition, another limitation of some studies is being task-dependent and as such not allowing for adaption in different interaction scenarios [135, 58, 96]. A promising approach seen is to use human social signals to design the reward function of the RL model and avoid large state spaces [139, 183, 117].

Adapting one modal robot behavior to a human social signal

Chapter 4 puts forward a pilot study that explores part of the goal by looking into how one human modality can be assessed by the robot to have it adapt one of its behavioral modalities and as a consequence permitting the robot to have a socially intelligent interaction. A social use case where adaptation is crucial is group interaction. The robot uses the human proxemics through their distances and orientations to estimate the conversational roles they want to play in the group formed: active, bystander, or overhearer. Accordingly, the robot autonomously adapted its gaze behavior to portray its adaptation to the group and roles. The primary results showed how consequently the humans themselves also adapted to the adaptive robot by standing closer to it in comparison to a robot that did not adapt its behavior to the roles in the group formed. In addition, the users rated the adaptive robot higher in the perceived adaptability and sociability. This study allowed to investigate one of the sub-questions of the main research question: *How can a human behavior be used to allow the robot to adapt to varying interaction setups?* by using group interaction as a use case. Moreover, these primary findings showed how one modality behavior of the robot can in its turn influence the behavior of the users. To further investigate this as well as the impact of each modality when performed in a multi-modal behavior, Chapter 5 presents the second study.

Synthesizing & Investigating Multi-Modal Behaviors of Robot

In Chapter 5, an autonomous multi-modal behavioral solution made of proxemics, gaze, and gesture was proposed and implemented. First, proxemics was implemented as a social navigation to maintain a close social distance. Second, gaze was designed as gaze aversions relative the human for the functional purposes of turn-taking, turn-yielding, floor-holding, and joint attention. Third, gesture carried out as a series of functional gestures including symbolic, deictic, and beat gestures. To investigate their respective impact in interactions, a study with 105 participants was done and with five conditions, where one condition has a full multi-modal behavior made of all modalities and the rest have one different modality each time subtracted from the full multi-modal behavior. The key findings showed that proxemics has a significant impact on how close humans stand to the robot while interacting with it in addition to influencing the human to take the robot's suggestion of an object in the same physical environment. In addition, one of the key findings pointed out to the strong effect of gesture has on the human also performing gestural behaviors during greeting and ending of interaction and as such adapting and aligning to the robot's behavior. Further findings show the significant impact of gaze on the backchanneling social signals performed by the human during the interaction. Thus, this study showed how when performed in a

multi-modal behavior the robot's modalities differently impact and influence the social signals of the human users and the outcome of the interaction. The bidirectional impact between the multi-modal behavior of the robot and the social signals of the humans, which are also inherently multi-modal, was demonstrated and thus moving closer to a socially intelligent interaction. One more layer of adaptation and personalization is needed to achieve that. Additionally, this study attempted to answer another sub-question: *How does each modality impact the outcomes of an interaction specially when it is performed in a multi-modal manner by the robot and in which aspect of the interaction is each modality relevant and significant?* .

Simulation of HRI setup & RL Model for Adaptation of Multi-Modal Behavior

The final study in Chapter 6 aims at closing the loop by proposing a RL model for multi-modal behavior and a simulation setup of HRI, which was designed and implemented using the data-set collected from the study in Chapter 5. Human social signals are continuous in nature and there is a need to avoid losing information by discretizing them and creating large state-spaces. The proposed RL model uses a reward function that takes into account the human's social signals and decisions and the cost on the complexity of the robot's behavior. The latter, which is the cost on the degree of multi-modal behaviors, allowed the investigation of whether complexity in behaviors of the robot is required for the success of an interaction. The reward signal is then used to adapt the behavior of the robot which may be a combination of various modalities including proxemics, gaze, gesture, and emotional expressions. The RL algorithms ϵ -MAB, ϵ -UCB, and Q-Learning were implemented and trained on the simulated HRI. The findings showed that complex behaviors are not necessarily needed for a socially intelligent interaction but rather appropriate combinations of behaviors that elicit the wanted social influence on the users. One of the key findings showed the importance of proxemics and how crucial it is for the engagement in HRI. All of the most selected multi-modal behaviors included proxemics. A second key finding exhibited the close role of gaze and gesture in an interaction. With Q-Learning, the most selected behaviors were the two two-modality behaviors, proxemics with gaze and proxemics with gesture, and followed by the three-modality behavior of proxemics with gaze and gesture. This demonstrated that the use of gaze and gesture may be interchangeable depending on the user profile and situation being adapted to. Finally, this thesis presents a generalized adaptation model that leaves room for the robot to personalize to different users and interactions. Moreover, these two contributions allowed to answer the following sub-question: *How can the robot use human social signals during an interaction to adapt and personalize its own multi-modal behavior?* .

7.1.2 Multi-Modal Behaviors & Social Intelligence: Perspective

Literature from various fields and sciences, including Psychology, Neuroscience, Sociology, and many others, have shown how complex yet rich human behaviors are. So it is no surprise that they have been observed, studied, and analyzed to help understand the social intelligence found in humans across cultures and countries. Thus for a robot to communicate clearly, increase engagement, and be perceived as socially intelligent during an interaction with a human, it needs to be able to appropriately use its multi-modal behavior and adapt to the multi-modal social signals of its interactive partner. In the first study, the significance of proxemics was put forward. By evaluating the proxemics of the humans through their distances and orientation using a simple rule-based system, the robot was able to estimate the conversational role intentions of participants forming a group around it. This was the first use of one of the modalities of the human's social signals to adapt the robot's behavior. The robot displayed its estimation and understanding through gaze alone. By using one modality alone, the robot was able to convey the message to the humans it is interacting with. The findings from the first pilot study showed that the humans in their turn felt more comfortable standing closer to the robot that adapted its gaze behavior than the one that did not. All the bidirectional communication between the robot and the humans forming a group around it to relay the conversational role intentions took place non-verbally and as such showing the power of multi-modal behavior. This lead to question how would other modalities and different combinations of them influence the human behavior in its turn.

In the second study investigating multi-modal behaviors, the findings further revealed the significance of multi-modal behaviors of the robot. First, proxemics was shown to influence the human to take recommendations made by the robot in the physical environment in comparison to other modalities. Previous literature in HRI have focused on gaze as joint attention and gesture as pointing to have the robot suggest objects in the environment around it [92, 93]. However, this study showed that when gaze, gesture, and proxemics were compared within multi-modal behaviors, it was proxemics that influenced the user taking the water on the table offered by the robot. It may be because proxemics allows the robot to use its embodiment to show that it exists in the space physical space as the human it is interacting with. Second, gesture had the most impact on how humans greeted and ended the interaction with the robot. Already from beginning of the interaction, performing gesture significantly influenced people waving hello to the robot. By the end of the interaction, participants, who interacted with the robot performing multi-modal behavior without gestures, did not wave goodbye to the robot. This is important as it shows that gestures affect how humans align their behaviors with the robot from early on. Even those who waved hello to the robot despite it not waving did not wave goodbye at the end. First impressions performed through

nonverbal behaviors impact the rest of the interaction with the robot. Third, gaze behavior impacts the backchanneling done by the human the robot is interacting with. Backchanneling is an interesting behavioral outcome to look at as it is done unconsciously by the human. When talking to participants after the data collection, some commented that the robot lacking gaze behavior seemed unnatural. It is important to remind that the condition lacking gaze behavior had the robot continuously look straight at the human's face. As humans, we are inclined to avert our gaze to signify that one is thinking, holding the floor of the conversation, or even giving the speaking role and similarly we respond and wait for such cues. Fourth, while social dialogue was also evaluated in comparison to the other behavioral modalities, no significant results for it were found. This further highlights the impact of non-verbal behavior and its ability to communicate and engage the user by only using the embodiment capabilities of the robot. Most literature in HRI have focused on parts of the interaction and on comparing one modality to another. This study introduced findings after looking at the whole interaction and performing multi-modal behaviors by removing a different modality at each condition. This offered a new perspective on understanding the effect of modalities making up the multi-modal behavior of the robot.

Finally, the proposed RL model trained on the simulation setup of HRI offered further interesting insights into multi-modal behaviors through the choices made by the learning agents. The importance of proxemics was again highlighted strongly. The effects of proxemics on the outcomes of an interaction can be directly observed through the distances held by the user and the decisions made as shown from the previous study. This impacted the designed reward function to favor proxemics. On the other hand, other robot modalities have more subtle effects on the human. Gaze and gesture seem to have an interchangeable role allowing for further personalization and adaptation to different user profiles. In addition, the study showed that multi-modal behaviors do not necessarily need to be complex and of high-degree to achieve a socially intelligent interaction but rather be made up of the appropriate modalities specially in the presence of proxemics. Throughout all three studies, proxemics had the biggest effect and value to achieve a socially intelligent interaction. Robots, unlike other virtual agents and computers, can navigate in the same physical space as the human. This has shown to influence the perceived conversational role one wants to have in an interaction as well as the decisions made throughout the rest of that interaction. Proxemics is a modality that needs to be further studied.

Moreover, social intelligence is a powerful tool humans possess that allows them to communicate effectively with the world and societies around them, manipulate situations, and even shift the way they are perceived to the outside world. The findings of the all three studies further put forward the significance of the robot's multi-modal behavior in achieving social

intelligence for the robot. It showed how potent the modalities of the robot’s multi-modal behavior are in shaping the behavior of the interactive partner and the environment around them. It allowed the robot to be perceived as not just a machine but as an embodied agent present in the same physical space and capable of manipulating it just like, to a certain extent, the human. There is still much work to be done on understanding the power of social intelligence and synthesizing it in robots and virtual agents even, but the finding of this thesis presented a foundation to build on socially intelligent robots.

7.1.3 Limitations & Future Work

While the thesis presents important findings, there are certain limitations to consider. First, there are several functional behaviors that make up the modalities of gaze, gesture, and proxemics beyond just the ones synthesized and studied in thesis. For instance, a gaze mechanism may also be designed to convey the cognitive process the robot and be perceived as thoughtful [74, 11] and there are various gesture mechanisms to be explored including iconic and beat gestures that play certain roles in a conversation [6]. Second, there are other modalities that may significantly influence an interaction such as haptics, which utilizes the sense of touch, and vocalics, which focuses on the tone and pitch of the voice. These modalities may impact how the user perceives a certain emotion or behavior. Future work would focus on exploring further modalities making up multi-modal behaviors. Third, emotional expressions were not sufficiently investigated in this thesis. Emotional expressions are themselves also multi-modal and cover a vast range of synthesized emotions from joy and excitement to fear and anger. It is important for future work to dive deeper into understanding the impact of emotional expressions on the perceived social intelligence of the robot. Fourth, the proposed RL model was only trained and analyzed using the simulation setup of HRI. Future work needs to investigate the suggested model in real life interactions. It would be highly interesting to see whether the choices of multi-modal behaviors selected by the learning agents would be similar and how the adaptation would impact the outcomes of interaction.

7.2 Conclusion

As we look ahead, it is evident that robots will play an ever larger role in our world and influence various fields such as healthcare, retail, and even enter our homes. Central to their success will be socially intelligent interactions and the ability to adapt and personalize to humans with different personalities, cultures, and disabilities through its multi-modal behavior that allow it to be present in the same space as the user. Thus, it is important to build inclusive technologies shaping a better and brighter future.

References

- [1] Abele, A. (1986). Functions of gaze in social interaction: Communication and monitoring. *Journal of Nonverbal Behavior*, 10(2):83–101.
- [2] Admoni, H., Datsikas, C., and Scassellati, B. (2014). Speech and gaze conflicts in collaborative human-robot interactions. In *Proceedings of Annual Meeting of the Cognitive Science Society (CogSci '14)*, volume 36, pages 104 – 109.
- [3] Admoni, H. and Scassellati, B. (2014). Data-driven model of nonverbal behavior for socially assistive human-robot interactions. In *Proceedings of the 16th international conference on multimodal interaction*, pages 196–199.
- [4] Admoni, H. and Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *J. Hum.-Robot Interact.*, 6(1):25–63.
- [5] Akalin, N. and Loutfi, A. (2021). Reinforcement learning approaches in social robotics.
- [6] Akbiyik, S., Karaduman, A., Goksun, T., and Chatterjee, A. (2018). The relationship between co-speech gesture production and macrolinguistic discourse abilities in people with focal brain injury. *Neuropsychologia*, 117.
- [7] Allport, G. (1959). *The Historical Background of Modern Social Psychology*.
- [8] Aly, A. and Tapus, A. (2013). A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 325–332.
- [9] Ammi, M., Demulier, V., Caillou, S., Gaffary, Y., Tsalamal, Y., Martin, J.-C., and Tapus, A. (2015). Haptic human-robot affective interaction in a handshaking social protocol. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, pages 263–270.
- [10] Andrist, S., Pejsa, T., Mutlu, B., and Gleicher, M. (2012). Designing effective gaze mechanisms for virtual agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 705–714, New York, NY, USA. Association for Computing Machinery.
- [11] Andrist, S., Tan, X. Z., Gleicher, M., and Mutlu, B. (2014). Conversational gaze aversion for humanlike robots. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '14, page 25–32, New York, NY, USA. Association for Computing Machinery.
- [12] Anzalone, S. M., Boucenna, S., Ivaldi, S., and Chetouani, M. (2015). Evaluating the engagement with social robots. *International Journal of Social Robotics*, 7(4):465–478.

- [13] Anzalone, S. M., Varni, G., Ivaldi, S., and Chetouani, M. (2017). Automated prediction of extraversion during human–humanoid interaction. *International Journal of Social Robotics*, 9(3):385–399.
- [14] Aoki, P. M., Szymanski, M. H., Plurkowski, L. D., Thornton, J. D., Woodruff, A., and Yi, W. (2006). Where’s the "party" in "multi-party"?: analyzing the structure of small-group sociable talk. In *CSCW*, pages 393–402. ACM.
- [15] Argyle, M. and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge University Press.
- [16] Argyle, M. and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 28:289–304.
- [17] Bagheri, E., Roesler, O., Cao, H.-L., and Vanderborght, B. (2020). A reinforcement learning based cognitive empathy framework for social robots. *International Journal of Social Robotics*.
- [18] Barnes, M. L. and Sternberg, R. J. (1989). Social intelligence and decoding of nonverbal cues. *Intelligence*, 13(3):263–287.
- [19] Biancardi, B., Dermouche, S., and Pelachaud, C. (2021). Adaptation mechanisms in human-agent interaction: Effects on user’s impressions and engagement. *Frontiers in Computer Science*.
- [20] Birnbaum, G. E., Mizrahi, M., Hoffman, G., Reis, H. T., Finkel, E. J., and Sass, O. (2016). Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 165–172.
- [21] Bolstad, C. A. (2001). Situation awareness: Does it change with age? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45(4):272–276.
- [22] Boucher, J.-D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., Dominey, P., and Ventre-Dominey, J. (2012a). I reach faster when i see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in Neurorobotics*, 6:3.
- [23] Boucher, J.-D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., Dominey, P., and Ventre-Dominey, J. (2012b). I reach faster when i see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in Neurorobotics*, 6:3.
- [24] Branigan, H. P., Pickering, M. J., Pearson, J., and McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, 42(9):2355 – 2368. How people talk to Robots and Computers.
- [25] Breazeal, C. L., Kidd, C. D., Thomaz, A. L., Hoffman, G., and Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pages 383–388.
- [26] Broadbent, E., Jayawardena, C., Kerse, N., Stafford, R. Q., and MacDonald, B. A. (2011). Human-robot interaction research to improve quality of life in elder care—an approach and issues. In *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.
- [27] Brockner, J. and Higgins, E. (2001). Regulatory focus theory: Implications for the study of emotions at work. *Organizational Behavior and Human Decision Processes*, 86:35–66.

- [28] Broekens, J. (2007). Emotion and reinforcement: Affective facial expressions facilitate robot learning. In Huang, T. S., Nijholt, A., Pantic, M., and Pentland, A., editors, *Artificial Intelligence for Human Computing*, pages 113–132, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [29] Brooks, R. A. (1999). *Cambrian intelligence: The early history of the new AI*. MIT press.
- [30] Brooks, R. A., Breazeal, C., Marjanović, M., Scassellati, B., and Williamson, M. M. (1998). The cog project: Building a humanoid robot. In *International workshop on computation for metaphors, analogy, and agents*, pages 52–87. Springer.
- [31] Brown, W. M. and Moore, C. (2002). Smile asymmetries and reputation as reliable indicators of likelihood to cooperate: An evolutionary analysis. In *11; 3*. Nova Science Publishers.
- [32] Campanella, S. and Belin, P. (2007). Integrating face and voice in person perception. *Trends in cognitive sciences*, 11(12):535–543.
- [33] Cappella, J. N. and Greene, J. O. (1982). A discrepancy-arousal explanation of mutual influence in expressive behavior for adult and infant-adult interaction. *Communications Monographs*, 49(2):89–114.
- [34] Chao, C. and Thomaz, A. L. (2013). Controlling social dynamics with a parametrized model of floor regulation. *Journal of Human-Robot Interaction*, 2(1):4–29.
- [35] Chartrand, T. and Bargh, J. (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of personality and social psychology*, 76(6):893–910.
- [36] Chiu, C.-y., Hong, Y.-y., and Krauss, R. M. (1995). Gaze direction and fluency in conversational speech. *Unpublished manuscript*.
- [37] Cialdini, R. and Goldstein, N. (2004). Social influence: Compliance and conformity. *Annual review of psychology*, 55:591–621.
- [38] Clark, H. H. and Krych, M. A. (2004). Speaking while monitoring addressees for understanding.
- [39] Côté, S., Hideg, I., and Van Kleef, G. A. (2013). The consequences of faking anger in negotiations. *Journal of Experimental Social Psychology*, 49(3):453–463.
- [40] Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical transactions of the royal society B: Biological sciences*, 362(1480):679–704.
- [41] Dautenhahn, K., Walters, M., Woods, S., Koay, K., Nehaniv, C., Sisbot, E., Alami, R., and Siméon, T. (2006). How may i serve you? a robot companion approaching a seated person in a helping context. volume 2006, pages 172–179.
- [42] De Melo, C. M., Carnevale, P. J., Read, S. J., and Gratch, J. (2014a). Reading people’s minds from emotion expressions in interdependent decision making. *Journal of personality and social psychology*, 106(1):73.
- [43] De Melo, C. M., Carnevale, P. J., Read, S. J., and Gratch, J. (2014b). Reading people’s minds from emotion expressions in interdependent decision making. *Journal of personality and social psychology*, 106(1):73.

- [44] de Melo, C. M. and Terada, K. (2019a). Cooperation with autonomous machines through culture and emotion. *PLOS ONE*, 14(11):1–12.
- [45] de Melo, C. M. and Terada, K. (2019b). Cooperation with autonomous machines through culture and emotion. *PLOS ONE*, 14(11):1–12.
- [46] Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-georges, C., Viaux, S., and Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3:349–365.
- [47] Dolinski, D., Nawrat, M., and Iza, R. (2001). Dialogue involvement as a social influence technique. *Personality and Social Psychology Bulletin*, 27:1395–1406.
- [48] Ekman, P. and Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1).
- [49] El Kaliouby, R., Picard, R., and Baron-Cohen, S. (2006). Affective computing and autism. *Annals of the New York Academy of Sciences*, 1093(1):228–248.
- [50] Ernst, M. O. and Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, 8(4):162–169.
- [51] Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., and Kreifelts, B. (2013). Functional responses and structural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *Neuroimage*, 76:45–56.
- [52] Fabiansson, E. C. and Denson, T. F. (2012). The effects of intrapersonal anger and its regulation in economic bargaining. *PLOS ONE*, 7(12):1–10.
- [53] Finn, C., Abbeel, P., and Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *CoRR*, abs/1703.03400.
- [54] Fiore, S. M., Wiltshire, T. J., Lobato, E. J. C., Jentsch, F. G., Huang, W. H., and Axelrod, B. (2013). Toward understanding social cues and signals in human – robot interaction : effects of robot gaze and proxemic behavior. 4(November):1–15.
- [55] Fiske, S. T., Cuddy, A. J., and Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences*, 11(2):77–83.
- [56] Fong, T., Thorpe, C., and Baur, C. (2003). Collaboration, dialogue, human-robot interaction. In Jarvis, R. A. and Zelinsky, A., editors, *Robotics Research*, pages 255–266, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [57] Ford, M. E. and Tisak, M. S. (1983). A further search for social intelligence. *Journal of Educational Psychology*, 75(2):196.
- [58] Gao, Y., Sibirtseva, E., Castellano, G., and Kragic, D. (2019). Fast adaptation with meta-reinforcement learning for trust modelling in human-robot interaction. *CoRR*, abs/1908.04087.
- [59] Gillet, S., Cumbal, R., Pereira, A., Lopes, J., Engwall, O., and Leite, I. (2021). Robot gaze can mediate participation imbalance in groups with different skill levels. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, HRI '21*, page 303–311, New York, NY, USA. Association for Computing Machinery.
- [60] Goffman, E. (1963). *Behavior in public places: notes on the social organization of gatherings*. Social theory. Free Press of Glencoe.

- [61] Goffman, E. (1967). *Interaction ritual: essays on face-to-face behavior*. Anchor books. Anchor Books.
- [62] Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological assessment*, 4(1):26.
- [63] Gratch, J. and de Melo, C. M. (2019). *Inferring Intentions from Emotion Expressions in Social Decision Making*, pages 141–160. Springer International Publishing, Cham.
- [64] Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.
- [65] Hall, E. and for the Anthropology of Visual Communication, S. (1974). *Handbook for proxemic research*. Studies in the anthropology of visual communication. Society for the Anthropology of Visual Communication.
- [66] Hall, E. and of Congress), C. P. C. L. (1966). *The Hidden Dimension*. Anchor books. Doubleday.
- [67] Hall, E. T. (1959). *The silent language / Edward Hall*. Doubleday Garden City, N.Y.
- [68] Hall, E. T. (1963). A system for the notation of proxemic behavior1. *American Anthropologist*, 65(5):1003–1026.
- [69] Ham, J., Bokhorst, R., Cuijpers, R., van der Pol, D., and Cabibihan, J.-J. (2011). Making Robots Persuasive: The Influence of Combining Persuasive Strategies (Gazing and Gestures) by a Storytelling Robot on Its Persuasive Power. In *Research on Education and Media*, volume 9, pages 71–83.
- [70] Handke, L. and Barthauer, L. (2019). *Heider (1958): The Psychology of Interpersonal Relations*, pages 259–262. Springer Fachmedien Wiesbaden, Wiesbaden.
- [71] Häring, M., Bee, N., and André, E. (2011). Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *2011 RO-MAN*, pages 204–209. IEEE.
- [72] Hasson, U. and Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693):20150366.
- [73] Heerink, M., Krose, B., Evers, V., and Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *I. J. Social Robotics*, 2:361–375.
- [74] Heylen, D. (2006). Head gestures, gaze and the principles of conversational structure. *International Journal of Humanoid Robotics*, 3(03):241–267.
- [75] Higgins, E. T. et al. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in experimental social psychology*, 30:1–46.
- [76] Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, 16(2):107–117.
- [77] Hornecker, E. (2008). “i don’t understand it either, but it is cool”-visitor interactions with a multi-touch table in a museum. In *2008 3rd IEEE international workshop on horizontal interactive human computer systems*, pages 113–120. IEEE.

- [78] Huang, C.-M. and Mutlu, B. (2012). Generating Effective Social Behaviors for Robots. *Proceedings of the 7th ACM/IEEE Conference on Human-Robot Interaction (HRI 2012)*, pages 25–32.
- [79] Huang, C.-M. and Mutlu, B. (2013). Modeling and evaluating narrative gestures for humanlike robots.
- [80] Immonen, A.-M. (2015). *Essays on emotional influences in consumer food choice: Understanding emotional intricacies in consumers' price vs. ethicality trade-off decisions, and perceptions of genetically modified food products*. PhD thesis.
- [81] Itti, L., Dhavale, N., and Pighin, F. (2003). Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation VI*, volume 5200, pages 64–78. International Society for Optics and Photonics.
- [82] Joosse, M., Poppe, R., Lohse, M., and Evers, V. (2014). Cultural differences in how an engagement-seeking robot should approach a group of people. In *Proceedings of the 5th ACM international conference on Collaboration across boundaries: Culture, Distance & Technology (CABS 2014)*, pages 121–130, United States. Association for Computing Machinery (ACM). eemcs-eprint-24983.
- [83] Kanda, T., Kamasima, M., Imai, M., Ono, T., Sakamoto, D., Ishiguro, H., and Anzai, Y. (2007). A humanoid robot that pretends to listen to route guidance from a human. *Autonomous Robots*, 22:87–100.
- [84] Kayukawa, Y., Takahashi, Y., Tsujimoto, T., Terada, K., and Inoue, H. (2017). Influence of emotional expression of real humanoid robot to human decision-making. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–6.
- [85] Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26(1):22–63.
- [86] Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Studies in Interactional Sociolinguistics. Cambridge University Press.
- [87] Kennedy, J., Baxter, P., and Belpaeme, T. (2015). The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15*, page 67–74, New York, NY, USA. Association for Computing Machinery.
- [88] Khamassi, M., Lallée, S., Enel, P., Procyk, E., and Dominey, P. F. (2011). Robot cognitive control with a neurophysiologically inspired reinforcement learning model. *Frontiers in neurorobotics*, 5:1.
- [89] Khamassi, M., Velentzas, G., Tsitsimis, T., and Tzafestas, C. (2018). Robot fast adaptation to changes in human engagement during simulated dynamic social interaction with active exploration in parameterized reinforcement learning. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):881–893.
- [90] Kim, A., Han, J., Jung, Y., and Lee, K. (2013). The effects of familiarity and robot gesture on user acceptance of information. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 159–160.

- [91] Kim, E. H., Kwak, S. S., Han, J., and Kwak, Y. K. (2009). Evaluation of the expressions of robotic emotions of the emotional robot, "mung". In *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication*, ICUIMC '09, page 362–365, New York, NY, USA. Association for Computing Machinery.
- [92] Kipp, M. and Martin, J.-C. (2009). Gesture and emotion: Can basic gestural form features discriminate emotions? *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–8.
- [93] Kirchner, N., Alempijevic, A., and Dissanayake, G. (2011). Nonverbal robot-group interaction using an imitated gaze cue. In *Proceedings of the 6th International Conference on Human-Robot Interaction*, HRI '11, page 497–504, New York, NY, USA. Association for Computing Machinery.
- [94] Kleef, G. A. V. (2009). How emotions regulate social life: The emotions as social information (easi) model. *Current Directions in Psychological Science*, 18(3):184–188.
- [95] Knox, W. B. (2012). Learning from human-generated reward.
- [96] Knox, W. B., Stone, P., and Breazeal, C. (2013). Training a robot via human feedback: A case study. In Herrmann, G., Pearson, M. J., Lenz, A., Bremner, P., Spiers, A., and Leonards, U., editors, *Social Robotics*, pages 460–470, Cham. Springer International Publishing.
- [97] Kong, A. P.-H., Law, S. P., Kwan, C., Lai, C., and Lam, V. (2015). A coding system with independent annotations of gesture forms and functions during verbal communication: Development of a database of speech and gesture (dosage). *Journal of Nonverbal Behavior*, 39.
- [98] Kosmitzki, C. and John, O. P. (1993). The implicit use of explicit conceptions of social intelligence. *Personality and individual differences*, 15(1):11–23.
- [99] Kruse, T., Kirsch, A., Sisbot, E. A., and Alami, R. (2010). Exploiting human cooperation in human-centered robot navigation. In *RO-MAN*, pages 192–197. IEEE.
- [100] Kubota, A. and Riek, L. D. (2022). Methods for robot behavior adaptation for cognitive neurorehabilitation. *Annual Review of Control, Robotics, and Autonomous Systems*, 5(1):null.
- [101] Kucherenko, T. (2018). Data driven non-verbal behavior generation for humanoid robots. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, ICMI '18, page 520–523, New York, NY, USA. Association for Computing Machinery.
- [102] Lee, J. and Marsella, S. C. (2006). Nonverbal Behavior Generator for Embodied Conversational Agents. In *6th International Conference on Intelligent Virtual Agents*, Marina del Rey, CA.
- [103] Lee, J. and Marsella, S. C. (2011). Modeling Side Participants and Bystanders: The Importance of Being a Laugh Track. In *Proceedings of the 11th Conference on Intelligent Virtual Agents*, Reykjavik, Iceland.
- [104] Lee, M. K., Forlizzi, J., Kiesler, S., Rybski, P., Antanitis, J., and Savetsila, S. (2012). Personalization in hri: A longitudinal field experiment. pages 319–326.

- [105] Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., and Paiva, A. (2011). Modelling empathy in social robotic companions. In *International conference on user modeling, adaptation, and personalization*, pages 135–147. Springer.
- [106] Lighthart, M. and Truong, K. P. (2015). Selecting the right robot: Influence of user attitude, robot sociability and embodiment on user preferences. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pages 682–687.
- [107] Lindblom, B. (1990). Explaining phonetic variation: A sketch of the h&h theory. In *Speech production and speech modelling*, pages 403–439. Springer.
- [108] Liu, C., Conn, K., Sarkar, N., and Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE transactions on robotics*, 24(4):883–896.
- [109] Liu, C., Ishi, C. T., Ishiguro, H., and Hagita, N. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '12, page 285–292, New York, NY, USA. Association for Computing Machinery.
- [110] Lucas, G. M., Boberg, J., Traum, D., Artstein, R., Gratch, J., Gainer, A., Johnson, E., Leuski, A., and Nakano, M. (2018). Getting to know each other: The role of social dialogue in recovery from errors in social robots. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '18, page 344–351, New York, NY, USA. Association for Computing Machinery.
- [111] Marshall, P., Rogers, Y., and Pantidi, N. (2011). Using f-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, CSCW '11, pages 445–454, New York, NY, USA. ACM.
- [112] McDermott, R. P. and Roth, D. R. (1978). The social organization of behavior: Interactional approaches. *Annual Review of Anthropology*, 7(1):321–345.
- [113] McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, Chicago.
- [114] Mead, R., Atrash, A., and Mataric, M. J. (2013). Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction. *I. J. Social Robotics*, 5(3):367–378.
- [115] Mead, R. A., Atrash, A., Kaszubski, E., Clair, A. S., Greczek, J., Clabaugh, C., Kohan, B., and Mataric, M. J. (2014). Building blocks of social intelligence: Enabling autonomy for socially intelligent and assistive robots. In *AAAI Fall Symposia*. AAAI Press.
- [116] Mehrabian, A. (1972). *Nonverbal Communication*. Aldine Publishing Company.
- [117] Mitsunaga, N., Smith, C., K, T., Ishiguro, H., and Hagita, N. (2008). Adapting robot behavior for human–robot interaction. *Robotics, IEEE Transactions on*, page 916.
- [118] Moerland, T. M., Broekens, J., and Jonker, C. M. (2018). Emotion in reinforcement learning agents and robots: a survey. *Machine Learning*, 107(2):443–480.
- [119] Moerland, T. M., Broekens, J., and Jonker, C. M. (2020). Model-based reinforcement learning: A survey. *arXiv preprint arXiv:2006.16712*.

- [120] Mol, L., Krahmer, E., and Swerts, M. (2009). Alignment in iconic gestures: Does it make sense? In Theobald, B.-J. and Harvey, R., editors, *Proceedings of the eight International Conference on Auditory-Visual Speech Processing (AVSP 2009)*, pages 3–8. School of Computing Sciences. Alignment in Iconic Gestures: Does it make sense?
- [121] Moon, A. J., Troniak, D. M., Gleeson, B., Pan, M. K., Zheng, M., Blumer, B. A., MacLean, K., and Crof, E. A. (2014). Meet me where i’m gazing: How shared attention gaze affects human-robot handover timing. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 334–341.
- [122] Moors, A., Ellsworth, P. C., Scherer, K. R., and Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, 5(2):119–124.
- [123] Morency, L.-P., de Kok, I., and Gratch, J. (2010). A probabilistic multimodal approach for predicting listener backchannels. *Autonomous Agents and Multi-Agent Systems*, 20(1):70–84.
- [124] Mussel, P., Göritz, A. S., and Hewig, J. (2013). The value of a smile: Facial expression affects ultimatum-game responses. *Judgment and Decision Making*, 8(3).
- [125] Mutlu, B., Kanda, T., Forlizzi, J., Hodgins, J., and Ishiguro, H. (2012). Conversational gaze mechanisms for humanlike robots. *ACM Transactions on Interactive Intelligent Systems*, 1:12.
- [126] Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., and Hagita, N. (2009a). Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI ’09*, page 61–68, New York, NY, USA. Association for Computing Machinery.
- [127] Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., and Hagita, N. (2009b). Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI ’09*, pages 61–68, New York, NY, USA. ACM.
- [128] Niculescu, A., van Dijk, B., Nijholt, A., Li, H., and See, S. L. (2013). Making social robots more attractive: the effects of voice pitch, humor and empathy. *International journal of social robotics*, 5(2):171–191.
- [129] Paradedda, R., Ferreira, M. J., Oliveira, R., Martinho, C., and Paiva, A. (2019). *What Makes a Good Robotic Advisor? The Role of Assertiveness in Human-Robot Interaction*, pages 144–154.
- [130] Park, E., Kong, H., Lim, H.-t., Lee, J., You, S., and del Pobil, A. P. (2011). The effect of robot’s behavior vs. appearance on communication with humans. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 219–220.
- [131] Pejisa, T., Gleicher, M., and Mutlu, B. (2017). Who, me? how virtual agents can shape conversational footing in virtual reality. In Beskow, J., Peters, C., Castellano, G., O’Sullivan, C., Leite, I., and Kopp, S., editors, *Intelligent Virtual Agents*, pages 347–359, Cham. Springer International Publishing.
- [132] Peters, R., Broekens, J., Li, K., and Neerincx, M. A. (2019). Robots expressing dominance: Effects of behaviours and modulation. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–7.

- [133] Peters, R., Broekens, J., and Neerincx, M. A. (2017). Robots educate in style: The effect of context and non-verbal behaviour on children’s perceptions of warmth and competence. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 449–455.
- [134] Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2):169–190.
- [135] Qureshi, A. H., Nakamura, Y., Yoshikawa, Y., and Ishiguro, H. (2017). Robot gains social intelligence through multimodal deep reinforcement learning. *CoRR*, abs/1702.07492.
- [136] Ramachandran, A., Litoiu, A., and Scassellati, B. (2016). Shaping productive help-seeking behavior during robot-child tutoring interactions. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 247–254. IEEE.
- [137] Reed, L. I., DeScioli, P., and Pinker, S. A. (2014). The commitment function of angry facial expressions. *Psychological Science*, 25(8):1511–1517. PMID: 24898726.
- [138] Reed, L. I., Zeglen, K. N., and Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous prisoner’s dilemma game. *Evolution and Human Behavior*, 33(3):200–209.
- [139] Ritschel, H. (2018). Socially-aware reinforcement learning for personalized human-robot interaction. In André, E., Koenig, S., Dastani, M., and Sukthankar, G., editors, *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018, Stockholm, Sweden, July 10-15, 2018*, pages 1775–1777. International Foundation for Autonomous Agents and Multiagent Systems Richland, SC, USA / ACM.
- [140] Ritschel, H., Kiderle, T., Weber, K., Lingenfeller, F., Baur, T., and André, E. (2020). Multimodal joke generation and paralinguistic personalization for a socially-aware robot. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 278–290. Springer.
- [141] Rudovic, O., Zhang, M., Schuller, B., and Picard, R. (2019). Multi-modal active learning from human data: A deep reinforcement learning approach. In *2019 International Conference on Multimodal Interaction*, pages 6–15.
- [142] Saad, E., Neerincx, M., and Hindriks, K. (2019). Welcoming robot behaviors for drawing attention. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 368–368, United States. IEEE. Video Abstract; 14th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2019 ; Conference date: 11-03-2019 Through 14-03-2019.
- [143] Salem, M., Rohlfing, K., Kopp, S., and Joubin, F. (2011). A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pages 247–252.
- [144] Sandstrom, G. M. and Dunn, E. W. (2014). Social interactions and well-being: The surprising power of weak ties. *Personality and Social Psychology Bulletin*, 40(7):910–922. PMID: 24769739.
- [145] Saunderson, S. and Nejat, G. (2019). How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, 11(4):575–608.

- [146] Saunderson, S. and Nejat, G. (2020). Investigating strategies for robot persuasion in social human-robot interaction. *IEEE Transactions on Cybernetics*.
- [147] Schegloff, E. A. (1987). Analyzing single episodes of interaction: an exercise in conversation analysis. *Social Psychology Quarterly*, 50(2):101–114.
- [148] Schegloff, E. A. (1998). Body torque. *Social Research*, 65(5):536–596.
- [149] Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., Pozzo, T., and Sandini, G. (2012). Measuring human-robot interaction through motor resonance. *International Journal of Social Robotics*.
- [150] Sharkey, A. (2014). Robots and human dignity: a consideration of the effects of robot care on the dignity of older people. *Ethics and Information Technology*, 16(1):63–75.
- [151] Shi, C., Shimada, M., Kanda, T., Ishiguro, H., and Hagita, N. (2011). Spatial formation model for initiating conversation. In *Spatial Formation Model for Initiating Conversation, Robotics: Science and Systems*.
- [152] Sidner, C. L., Lee, C., Morency, L.-P., and Forlines, C. (2006). The effect of head-nod recognition in human-robot conversation. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, HRI '06*, page 290–296, New York, NY, USA. Association for Computing Machinery.
- [153] Silvera, D., Martinussen, M., and Dahl, T. I. (2001). The tromsø social intelligence scale, a self-report measure of social intelligence. *Scandinavian journal of psychology*, 42(4):313–319.
- [154] Sinaceur, M., Adam, H., Van Kleef, G. A., and Galinsky, A. D. (2013). The advantages of being unpredictable: How emotional inconsistency extracts concessions in negotiation. *Journal of Experimental Social Psychology*, 49(3):498–508.
- [155] Sinaceur, M. and Tiedens, L. Z. (2006). Get mad and get more than even: When and why anger expression is effective in negotiations. *Journal of Experimental Social Psychology*, 42(3):314–322.
- [156] Sisbot, E. A., Marin-Urias, L. F., Alami, R., and Simeon, T. (2007). A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 23(5):874–883.
- [157] Skantze, G., Hjalmarsson, A., and Oertel, C. (2014). Turn-taking, feedback and joint attention in situated human-robot interaction. *Speech Commun.*, 65:50–66.
- [158] Smith, C. A., Haynes, K. N., Lazarus, R. S., and Pope, L. K. (1993). In search of the "hot" cognitions: Attributions, appraisals, and their relation to emotion. *Journal of personality and social psychology*, 65(5):916.
- [159] Srinivasan, V., Bethel, C. L., and Murphy, R. R. (2014). Evaluation of head gaze loosely synchronized with real-time synthetic speech for social robots. *IEEE Transactions on Human-Machine Systems*, 44(6):767–778.
- [160] Stanton, C. and Stevens, C. J. (2014). Robot pressure: The impact of robot eye gaze and lifelike bodily movements upon decision-making and trust. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pages 330–339.

- [161] Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, USA.
- [162] Sutton, R. S., Maei, H. R., Precup, D., Bhatnagar, S., Silver, D., Szepesvári, C., and Wiewiora, E. (2009). Fast gradient-descent methods for temporal-difference learning with linear function approximation. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 993–1000.
- [163] Sutton, R. S., McAllester, D., Singh, S., and Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. In *Proceedings of the 12th International Conference on Neural Information Processing Systems, NIPS'99*, page 1057–1063, Cambridge, MA, USA. MIT Press.
- [164] Szafir, D. and Mutlu, B. (2012). Pay attention! In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, page 11, New York, New York, USA. ACM Press.
- [165] Takahashi, Y., Kayukawa, Y., Terada, K., and Inoue, H. (2021). Emotional expressions of real humanoid robots and their influence on human decision-making in a finite iterated prisoner’s dilemma game. *International Journal of Social Robotics*, pages 1–10.
- [166] Tanevska, A., Rea, F., Sandini, G., Cañamero, L., and Sciutti, A. (2020). A socially adaptable framework for human-robot interaction. *Frontiers in Robotics and AI*, 7:121.
- [167] Tatarian, K., Chamoux, M., Pandey, A. K., and Chetouani, M. (2021a). Robot gaze behavior and proxemics to coordinate conversational roles in group interactions. In *2021 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN)*, pages 1297–1304.
- [168] Tatarian, K., Stower, R., Rudaz, D., Chamoux, M., Kappas, A., and Chetouani, M. (2021b). How does modality matter? investigating the synthesis and effects of multi-modal robot behavior on social intelligence. *International Journal of Social Robotics*, pages 1–19.
- [169] Thorndike, E. L. (1920). Intelligence and its use. *Harper’s Magazine*, (140):227–235.
- [170] Todi, K., Bailly, G., Leiva, L., and Oulasvirta, A. (2021). Adapting user interfaces with model-based reinforcement learning. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13.
- [171] Turkstra, L., Ciccia, A., and Seaton, C. (2003). Interactive behaviors in adolescent conversation dyads. *Language Speech and Hearing Services in Schools*, 34:117–127.
- [172] van Baaren, R. B., Holland, R. W., Kawakami, K., and van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*, 15(1):71–74. PMID: 14717835.
- [173] van Kleef, G. A. and Côté, S. (2018). Emotional dynamics in conflict and negotiation: Individual, dyadic, and group processes. *Annual Review of Organizational Psychology and Organizational Behavior*, 5(1):437–464.
- [174] Van Maris, A., Sutherland, A., Mazel, A., Dogramadzi, S., Zook, N., Studley, M., Winfield, A., and Caleb-Solly, P. (2020). The impact of affective verbal expressions in social robots. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 508–510.

- [175] Vanderdonckt, J., Bouzit, S., Calvary, G., and Chêne, D. (2019). Exploring a design space of graphical adaptive menus: normal vs. small screens. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(1):1–40.
- [176] Vázquez, M., Carter, E. J., McDorman, B., Forlizzi, J., Steinfeld, A., and Hudson, S. E. (2017). Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI '17*, page 42–52, New York, NY, USA. Association for Computing Machinery.
- [177] Vermorel, J. and Mohri, M. (2005). Multi-armed bandit algorithms and empirical evaluation. In *European conference on machine learning*, pages 437–448. Springer.
- [178] Vertegaal, R., Slagter, R., van der Veer, G., and Nijholt, A. (2001). Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01*, page 301–308, New York, NY, USA. Association for Computing Machinery.
- [179] Vinciarelli, A. and Pentland, A. S. (2015). New social signals in a new interaction world: the next frontier for social signal processing. *IEEE Systems, Man, and Cybernetics Magazine*, 1(2):10–17.
- [180] Wagner, J., Lingensfelder, F., Baur, T., Damian, I., Kistler, F., and André, E. (2013). The social signal interpretation (ssi) framework: Multimodal signal processing and recognition in real-time. In *Proceedings of the 21st ACM International Conference on Multimedia, MM '13*, page 831–834, New York, NY, USA. Association for Computing Machinery.
- [181] Wallkötter, S., Stower, R., Kappas, A., and Castellano, G. (2020). A robot by any other frame: Framing and behaviour influence mind perception in virtual but not real-world environments. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, HRI '20*, page 609–618, New York, NY, USA. Association for Computing Machinery.
- [182] Wang, Y., Lucas, G., Khooshabeh, P., De Melo, C., and Gratch, J. (2015). Effects of emotional expressions on persuasion. *Social Influence*, 10(4):236–249.
- [183] Weber, K., Ritschel, H., Aslan, I., Lingensfelder, F., and André, E. (2018). How to shape the humor of a robot - social behavior adaptation based on reinforcement learning. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction, ICMI '18*, page 154–162, New York, NY, USA. Association for Computing Machinery.
- [184] Wiltshire, T. J., Lobato, E. J. C., Garcia, D. R., Fiore, S. M., Jentsch, F. G., Huang, W. H., and Axelrod, B. (2015). Effects of robotic social cues on interpersonal attributions and assessments of robot interaction behaviors. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59(1):801–805.
- [185] Wood, L. J., Zarak, A., Walters, M. L., Novanda, O., Robins, B., and Dautenhahn, K. (2017). The iterative development of the humanoid robot kaspar: an assistive robot for children with autism. In *International Conference on Social Robotics*, pages 53–63. Springer.
- [186] Xu, J., Broekens, J., Hindriks, K., and Neerincx, M. A. (2015). Mood contagion of robot body language in human robot interaction. *Autonomous Agents and Multi-Agent Systems*, 29(6):1216–1248.

- [187] Yohanan, S. and MacLean, K. E. (2012). The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *International Journal of Social Robotics*, 4(2):163–180.
- [188] Yonezawa, T., Yamazoe, H., Utsumi, A., and Abe, S. (2007). Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking. In *Proceedings of the 9th International Conference on Multimodal Interfaces, ICMI '07*, page 140–145, New York, NY, USA. Association for Computing Machinery.
- [189] Young, A. W. and Bruce, V. (2011). Understanding person perception. *British journal of psychology*, 102(4):959–974.
- [190] Zecca, M., Mizoguchi, Y., Endo, K., Iida, F., Kawabata, Y., Endo, N., Itoh, K., and Takanishi, A. (2009). Whole body emotion expressions for kobian humanoid robot — preliminary experiments with different emotional patterns —. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 381–386.
- [191] Zhang, H. and Yu, T. (2020). *Taxonomy of Reinforcement Learning Algorithms*, pages 125–133. Springer Singapore, Singapore.